

AD-A279 175



**DEFENSE**

**MANPOWER DATA CENTER**

**POPULATION PROPENSITY  
MEASUREMENT MODEL**

**Final Analysis Report**

***Market Research and Analysis Branch***

**94-14530**



**94 5 13 092**

**1600 WILSON BOULEVARD ARLINGTON, VIRGINIA 22209**

**DTIC QUALITY INSPECTED 1**

# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December, 1993	3. REPORT TYPE AND DATES COVERED Final	
4. TITLE AND SUBTITLE Population Propensity Measurement Model: Final Analysis Report			5. FUNDING NUMBERS MDA903-90-C-0195	
6. AUTHOR(S) Brice Stone, Kathryn Turner, Vincent Wiggins				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) RRC, Inc. 3833 Texas Ave, Suite 285 Bryan, TX 77802			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Manpower Data Center Attn: Jerome Lehnus 1600 N. Wilson Blvd., Suite 400 Arlington, VA 22209-2593			10. SPONSORING/MONITORING AGENCY REPORT NUMBER DMDC Report 93-012	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT <b>DISTRIBUTION STATEMENT A</b> Approved for public release; Distribution Unlimited		DTIC ELECTE MAY 16 1994 S B D		12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) The Propensity Measurement Model Final Report presents analyses of enlistment likelihood and estimates of high and low quality measures using data collected in the Youth Attitude Tracking Study. For the analyses, survey data were matched to military records to develop models for estimating both performance on military qualifying (AFQT) tests and on the probability of future enlistment. Report includes a review of literature on enlistment intentions, a discussion of methodology used and replication strategy for deriving enlistment behavior from survey data, and models for replicating the study from future data collections.				
14. SUBJECT TERMS Model, Enlistment Intention, AFQT, YATS			15. NUMBER OF PAGES 131	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR	

# **POPULATION PROPENSITY MEASUREMENT MODEL**

## **Final Analysis Report**

Contract No. MDA903-90-C-0195

### **SUBMITTED TO:**

**Defense Manpower Data Center**  
1600 Wilson Blvd.  
Arlington, Virginia 22209

### **SUBMITTED BY:**

**Dr. Brice M. Stone**  
**Ms. Kathryn L. Turner**  
and  
**Mr. Vincent L. Wiggins**

For Additional Copies of Report,  
Contact:

Defense Technical Information Center (DTIC)  
Cameron Station, Building 5  
Alexandria, VA 22304-6145

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By <i>per telecom</i>	
Distribution	
Availability Codes	
Dist	Avail and/or Special
<i>A-1</i>	

Statement A. Auth: DMDC  
(Mr. Tom Ulrich 703-696-6675)  
telecom,

5/16/94 CB

## TABLE OF CONTENTS

I. INTRODUCTION . . . . .	1
II. REVIEW OF YATS "PROPENSITY TO APPLY" LITERATURE . . . . .	3
Intentions to Join . . . . .	3
The Effect of Demographic Factors on the Intention to Join . . . . .	5
Negative Intentions to Join . . . . .	6
Male and Female Intentions to Join . . . . .	7
The Affect of Geodemographic Clusters on the Intention to Join . . . . .	8
III. YATS AND THE PROPENSITY TO APPLY . . . . .	11
Matching YATS and MEPS Data . . . . .	11
Sample Population . . . . .	14
IV. REPLICATION AND EXTENSION OF EARLIER INTENTION ANALYSIS . . . . .	17
Variable Description . . . . .	17
Estimation of the Application Equation . . . . .	18
Preliminary Extension of the Application Equation . . . . .	21
Including Males and Females in the Application Equation . . . . .	22
V. PPMM ESTIMATION . . . . .	27
Intensity of Intention on the Propensity to Apply . . . . .	28
Potential Biases in the Application Data . . . . .	28
Schooling and Non-Schooling Groups of Respondents . . . . .	29
Other Explanatory Variables from YATS . . . . .	31
The Importance of the Population Distribution . . . . .	33
Estimation Results for PPMM . . . . .	35
Predicting Application Rates Using PPMM . . . . .	40
VI. MULTIPLE DECISION MODELS . . . . .	43
Multinomial Logit Applied to Work, School, and Military . . . . .	45
Multinomial Logit Results . . . . .	47
VII. MODEL METHODOLOGY AND PROJECTION . . . . .	55

Respondent Selection . . . . .	55
Projection Model Equation . . . . .	56
Application of the Projection Model Equation . . . . .	59
Compute Fiscal Year Average Probabilities . . . . .	60
Adjust Fiscal Year Estimates . . . . .	61
Methodology Summary . . . . .	68
VIII. CONCLUSIONS AND RECOMMENDATIONS . . . . .	69
REFERENCES . . . . .	73
APPENDIX A. REVIEW OF ANALYSIS OF INTENTION INFORMATION . . . . .	77
APPENDIX B. RE-ESTIMATION OF THE QUALITY EQUATION . . . . .	89
APPENDIX C. CREATION OF VARIABLES FROM YATS DATA . . . . .	95
APPENDIX D. YATS QUESTIONS INCLUDED IN ANALYSIS . . . . .	103
APPENDIX E. HIGH QUALITY ENLISTMENT CONTRACTS MODEL . . . . .	113

## TABLE OF TABLES

<u>Table</u>	<u>Page</u>
1. Distribution of DIFFDT Using MEPS Date . . . . .	12
2. Distribution of SIGNDT . . . . .	13
3. SSANs Provided by Year of YATS . . . . .	15
4. SSANs Provided by Females by Year of YATS . . . . .	15
5. Variable Definitions . . . . .	19
6. Results of the Replication of Earlier Propensity Models . . . . .	20
7. Results Including New Intention Measure . . . . .	22
8. Males and Females (16 to 21) in the YATS . . . . .	23
9. Results for Male and Female Equation . . . . .	24
10. New Intention Variables . . . . .	28
11. New Variables for ASVAB-Taking and Recruiter Contact . . . . .	29
12. Percent Contacted Recruiter and Taken ASVAB . . . . .	30
13. Percent Applied by Recruiter and ASVAB Groups . . . . .	30
14. New Variables for Schooling . . . . .	31
15. Other Explanatory Variable Definitions . . . . .	32
16. Variation in Intention Over Time . . . . .	33
17. Variation in the Application Rate Over Time . . . . .	34
18. Population Variation Compared with Application Rates . . . . .	34
19. PPMM Estimates for Males, 16 to 18 Years of Age . . . . .	36
20. PPMM Estimates for Males, 19 to 21 Years of Age . . . . .	37
21. PPMM Estimates for Females, 16 to 18 Years of Age . . . . .	38
22. PPMM Estimates for Females, 19 to 21 Years of Age . . . . .	39
23. Predicted Application Rate for Males, 16 to 18 Years of Age . . . . .	41
24. Predicted Application Rate for Males, 19 to 21 Years of Age . . . . .	41
25. Predicted Application Rate for Females, 16 to 18 Years of Age . . . . .	42
26. Predicted Application Rate for Females, 19 to 21 Years of Age . . . . .	42
27. Dependent Variable - Multinomial Logit . . . . .	46
28. Military v. Work v. School . . . . .	46
29. Military v. Work v. School by Age . . . . .	47
30. Binary Variables for Multinomial Logit Estimation . . . . .	48
31. Young Males - Military/School - Work/School . . . . .	49
32. Young Males - Military/Work - School/Work . . . . .	50
33. Older Males - Military/School - Work/School . . . . .	51
34. Older Males - Military/Work - School/Work . . . . .	52
35. Young Females - Military/School - Work/School . . . . .	53
36. Young Females - Military/Work - School/Work . . . . .	54
37. YATS Respondents Applying Within Three Years of YATS . . . . .	56
38. New Variables in the Projection Model . . . . .	58
39. Projection Model Estimate Males and Females, 18 to 21 During Application FY . . . . .	59

40. Projected and Actual Annual Application Rates for Youth Age 17 to 21 . . . . .	65
41. Projected and Actual Annual Application Rates for Females Age 18 to 21 . . . . .	66
42. Summary of Model Methodology . . . . .	68
B1. Variable Definitions for Quality Equation . . . . .	92
B2. Quality Equation Re-estimation Results . . . . .	93
E1. Variable Definitions . . . . .	114
E2. Army Estimation Results . . . . .	117
E3. Navy Estimation Results . . . . .	118
E4. Air Force Estimation Results . . . . .	120
E5. Marine Corps Estimation Results . . . . .	121

## TABLE OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1. Comparison of weighted projections and actual application rates (from MEPCOM data) for 18 to 21 year old males from FY 1986 through 1991. . . . .	62
Figure 2. Comparison of weighted projections and actual application rates (from MEPCOM data) for 18 to 21 year old females from FY 1986 through 1991. . . . .	62
Figure 3. Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 17 to 21 year old males from FY 1986 through 1991. . . . .	65
Figure 4. Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 17 to 21 year old females from FY 1986 through 1991. . . . .	66
Figure 5. Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 18 to 21 year old females from FY 1986 through 1991. . . . .	67

## **I. INTRODUCTION**

The development of the Population Propensity Measurement Model (PPMM) had two objectives: (1) to design a model to forecast enlistments which contains attitudinal and behavioral data from Youth Attitude Tracking Study (YATS) questions which are deemed empirically relevant to the forecasting of enlistment rates combined with external factors and policy malleable data and (2) to develop a computer-based model which includes the specification of the forecast model(s) and all data used in the model development and enlistment projection. The need for a model or models to project the ease or difficulty of attaining future recruiting goals is apparent from a budget requirements and mission readiness perspective. The flow of personnel into the military is affected by numerous demographic and attitudinal factors.

The YATS is the primary source of demographic and attitudinal data for the Department of Defense (DoD). The combination of YATS data with military personnel data and personnel policy information provides an opportunity to analyze and estimate a propensity to enlist or take the Armed Forces Vocational Aptitude Battery (ASVAB) model(s) which can be used in assessing future enlistment or ASVAB taking rates or likelihoods.

Section II presents a review of the literature relevant for the design and development of PPMM using data from the YATS. Several key studies are presented in Section II which provide the basis for comparison of the PPMM design and estimation. Section III details the methodology by which Military Enlistment Processing Station (MEPS) applicant data can be used in conjunction with the YATS survey data to develop models which predict the propensity of YATS respondents to take the ASVAB or apply for military service. Section IV presents the results of replicating some of the earlier analysis performed using YATS data from an earlier time period. Comparisons are made between the model specification and estimation results of the earlier YATS studies and the present study's replication of those results using YATS data from more recent years. Section V presents the estimation results of the parameters for PPMM. In addition, Section VI provides analysis results using a multinomial logit approach to estimating the parameters of a multiple outcome decision model. Section VII presents the methodology

used to project application rates for future youth cohorts based on YATS responses and PPMM. Section VIII provides conclusions and recommendations for future research. Appendix E of the report contains an aggregate model which can be used to predict high quality enlistment contracts for each Service.

## II. REVIEW OF YATS "PROPENSITY TO APPLY" LITERATURE

Over the period 1982 to 1989, a series of studies were performed by Orvis, Gahart, and Hosek which explored the usefulness of intention information, primarily from the Youth Attitude Tracking Study (YATS), in predicting enlistment behavior. Appendix A provides a detailed summary of each of these articles. These studies also developed ways to identify the quality of individuals (ASVAB scores) who have not yet tested for entry into the Armed Services. Appendix B presents a complete discussion of the re-estimation of a quality equation with more recent YATS data. Also, the question of whether geodemographic information aids in the identification of high quality youths was investigated.

The YATS and MEPS data bases contain information on background characteristics, economic factors, educational experiences, and military interest indicators. Additional data bases which were used in the studies were the National Longitudinal Surveys of Youth Labor Market Experiences (NLS), the Cotterman data base, and A Clustering of Residential Neighborhoods (ACORN) information. Armed Forces Qualifying Test (AFQT) scores were obtained from Military Entrance Processing Reporting Systems (MEPRS) historical data files to document AFQT scores for YATS respondents who were tested.

### Intentions to Join

Orvis (1982) used two questions from the YATS survey to construct a measure of the respondent's intention to join the military. *Question 1* asked the respondent *What do you plan on doing in the next few years?* *Question 2* asked the respondent *How likely is it that you will join the military service?* *Question 1* is referred to as an *unaided mention* question since the respondent is not provided choices from which to select a response and, at that point in the interview, military enlistment has not been mentioned by the interviewer. Respondents are prompted for more than one response, up to a maximum of five responses. *Question 2* is referred to as an *intention question*. The respondent to *Question 2* is provided with four possible responses: definitely, probably, probably not, and definitely not. Orvis uses these two questions

to construct a measure of propensity for each respondent which reflects the propensity of the individual to join the Service. Four intention groups based on these two questions were formed:

*Group 1. Unaided mention, definite intention*

*Group 2. Unaided mention, probable intention*

*Group 3. No unaided mention, definite or probable intention, and*

*Group 4. No unaided mention, probably not or definitely not intention.*

Orvis (1982) begins by examining the characteristics of YATS respondents who eventually apply for military service. To examine these issues, distributions of the respondent's enlistment actions over time were computed and compared to the distribution of those taking the written test to enter military service. This initial analysis revealed that the number of enlistments and written exams increased throughout a 42 month period. This indicates the need for long-term follow-ups of studies modeling enlistment decisions. Additional analysis was performed that compared the characteristics of the respondents who took the written test to the population as a whole. In general, the persons taking the test tended to be younger, were less likely to be high school graduates, and were more likely to be from a minority background. Overall, the data suggest that enlistment intention measures are valid for both high and low quality applicants.

Analysis was then conducted that examined the relationship between intentions and enlistment. Groups were stratified by quality and propensity level. Orvis showed that the pattern of enlistment does not vary by quality of the applicant, and that the strongest propensities to enlist are associated with the highest stated intention levels. Enlistment rates ranged from 11% for those who expressed negative propensity to join military service (Group 4) to 53% for those expressing positive propensity (Groups 1 through 3).

Comparisons were also made that looked at the propensity to join the military in general versus expressed intent to join a specific Service. Questions Q510 through Q513 ask the respondent *How likely is it that you will be joining the Army/Air Force/Marine Corps/Navy?* The respondents may choose from four possible responses: definitely, probably, probably not, and definitely not. The results indicate that the Service-specific measures perform better than the general measure in terms of predicting enlistment behavior for a particular Service. As an example, 33 percent of the individuals with an unaided mention and definite intention to serve

in the Army actually enlisted in the military within one year. This compares to an 18% enlistment rate for individuals with an unaided mention and definitive intention to serve in the military in general.

### **The Effect of Demographic Factors on the Intention to Join**

Orvis (1984) builds upon previous work. Even though a strong relationship was found to exist between enlistments and intentions from Orvis (1982), it was thought that the intention information may be completely captured by other demographic factors. If this were the case, then intention information would not add to the predictive capability of the enlistment behavior model.

In order to test this proposition, enlistment rates were analyzed across the four intention groups. Comparisons were made with the negative propensity group. After controlling for a variety of demographic factors, the enlistment rate for the unaided mention-definite intention group was 36% higher than the enlistment rate of the negative propensity group (Group 4). Similarly, the increase in the enlistment rate was 23% higher for the unaided mention-probable intention group and 7% higher for the positive propensity-no unaided mention. A corresponding analysis that examined the percent of respondents taking the Armed Service Vocational Aptitude Battery (ASVAB) was conducted. The resulting increase in the enlistment rate was 41% for Group 1, 33% for Group 2, and 12% for Group 3 when compared to the negative propensity group. Thus, intention data was shown to add to the predictive capability of the enlistment model that includes demographic characteristics.

Given that intention data adds to the ability to predict enlistment decisions, the next step was to assess whether or not it could help in predicting first-term enlisted performance. The performance measures studied were attrition and promotion. These measures were regressed on the same demographic factors. No significant relationship was found to exist between enlistment intentions and promotion. The results also showed no significant difference between the attrition rates of the negative propensity and the positive propensity-no aided mention groups, the largest two cohorts.

Finally the work focused on the ability of the model to predict AFQT scores based on demographic characteristics. Similar to other work, the respondents' AFQT percentile scores were found to have a positive relationship with grade point average, the number of math courses completed, and father's education. AFQT scores were found to be lower for minorities and respondents from the South. The model was able to correctly classify about 70-75 percent of the respondents into either the upper or lower half of the population.

### **Negative Intentions to Join**

Earlier work by Orvis (1982) and Orvis (1984) indicated that those with negative intentions were an important source of recruits for the military. Orvis and Gahart (1985) began by examining and contrasting the similarities and differences between the negative and positive intention groups. If the two groups proved to be similar, then this would suggest that a common recruiting approach could attract individuals from both groups. If not, then recruiting strategy would have to be targeted to the specific groups.

The approach employed by Orvis and Gahart used those with negative intentions as a control group and performed regression analysis to explain differences in enlisting and testing rates across the cohorts in question. The enlistment and testing rates in the positive intention groups were compared to those in the negative intention group. After controlling for differences in background characteristics, the analysis revealed that persons in the first group (unaided mention and positive intentions), on average displayed enlistment rates 24% higher and testing rates 30% higher than the control group (negative intention). Respondents in the second group (no unaided mention and positive intentions) displayed enlistment rates 5% higher and test taking rates 9% higher than the negative intention group. Further analysis revealed that even though only 6% of the individuals with negative intentions enlisted in military service, they accounted for 46% of total enlistees. This is mainly attributed to the size of this group. This fact indicates that small increases in the enlistment rate for this group can account for a substantial increase in the number of enlistees. This indicated that the characteristics of enlistees in each group need to be analyzed.

Analysis of the positive and negative intention groups for respondents in high school was performed to address the issue of whether or not the model could distinguish enlistees from non-enlistees. Both long-term and short-term factors were identified to influence the enlistment decision in the two groups. In many cases the factors identified were the same. In general, those respondents who were of minority status, had taken fewer math courses, perceived difficulty finding work, had discussed enlisting in military service, and perceived the military as offering job security, displayed higher enlistment rates. This same analysis was also performed on non high school respondents. As before, the results with respect to the negative and positive intention groups were similar.

Next, the relationship between aggregate enlistment rates and aggregate intention levels in current and future periods was examined. A time-series cross-sectional analysis was performed for 17 geographical regions. These results indicate that there is a significant relationship between intention levels and concurrent high quality enlistments. Analysis was then performed to determine if there were any lagged effects of intention on enlistments. After controlling for current intention, the lagged intention measure was significant in three out of four analyses.

### **Male and Female Intentions to Join**

Since women were not included in YATS surveys until the Fall 1980, initial work by Orvis (1982 and 1984) and Orvis and Gahart (1985) only examined the relationship between survey enlistment intention measures and non prior service (NPS) male respondents' subsequent enlistment decisions. Orvis (1982) and Orvis and Gahart (1985) indicated that a long lag structure existed between stated intentions and the enlistment decision. Therefore, the Orvis (1986) study was intended to highlight the results of male respondents and to provide the same analysis for female respondents.

The measures used were the same intention and unaided mention groups as used in previous studies. The data show that approximately 87% of the females expressed negative intentions toward joining the military, compared to only about 68% of male respondents expressing negative intentions. An examination of the distributions of the number of enlistees

was performed across three groups: 1) positive intention and unaided mention 2) positive intention and no unaided mention and 3) negative intention. The distributions revealed that 18% of the females in Group 1 enlisted in the military, 4% in Group 2 enlisted, and 1% in Group 3 enlisted before the follow-up period. For male respondents 31% in Group 1 enlisted, 12% in Group 2 enlisted, and 6% in Group 3 enlisted. The pattern for males and females was found to be similar with respect to the percent of respondents taking the ASVAB. It is hypothesized that demand constraints on the types of jobs available, physical requirements, and social pressures tend to limit enlistment among women. Further analysis revealed, after controlling for differences in background characteristics, that the influence of positive intentions on enlistment testing rate was similar for both sexes. The results with respect to enlistments were also similar though not as strong as those with respect to intentions.

Overall, the data revealed that despite the low enlistment rate for females with negative intentions, this group accounted for 63% of actual females enlistees. Fourteen percent of female enlistees came from Group 1 and the remaining 23% from Group 2. This compared to 46% from Group 3, 19% from Group 1, and 35% from Group 2 for actual male enlistees (Orvis and Gahart, 1985). This analysis revealed that lumping intention data for males and females together tended to overstate female intentions to enlist.

### **The Effect of Geodemographic Clusters on the Intention to Join**

The work by Orvis, Gahart, and Hosek (1989) was intended to synthesize and extend the results of previous work and to examine the question of whether geodemographic clusters add to the predictive power of models of individual enlistment decision making. It also investigated whether the factors used in predicting geodemographic models vary across geodemographic groups. Finally, this report examines the relationship between ACORN (A Clustering of Residential Neighborhoods) information and micro models of enlistment behavior.

Since research up to this point had only validated the use of ZIP code level information, this paper begins by examining the relationship between estimates of ACORN information produced by the FIPS and ZIP code approaches. The results of this analysis showed that the FIPS and ZIP code measures produce similar results. Having determined this, the next effort

was to determine the relationship between geodemographic data bases and individual-level micro models. Comparisons between actual and predicted enlistment rates for the 44 ACORN clusters within each county were examined. Results of the analysis showed a highly significant relationship between observed and predicted enlistment rates. This indicates that the micro models accounted for most of the variation in enlistment rates among the ACORN clusters. Likewise, analysis was performed using another data base. This analysis confirmed the previous finding that the variables in the micro model do a good job of explaining the variation in enlistment rates across geodemographic clusters.

Finally, data were used to determine if ACORN information could improve the predictive capability of the micro models. Stepwise regression was performed that provided little evidence of a significant increase in explanatory power from the geodemographic information. In addition, three logistic regression equations were analyzed that used ACORN information, micro-model information, and both ACORN and micro-model information. The predicted probabilities from each model were ranked and grouped into five quintiles. The results of this analysis showed that even though ACORN information alone accounted for some variation in enlistment rates, the micro-model accounted for much more of the variation. Also, when the ACORN information was added to the micro-model, there was no statistical difference in the predicted enlistment rates.

### **III. YATS AND THE PROPENSITY TO APPLY**

The propensity of an individual to enlist in the military measures the likelihood of that individual, under specific conditions, to apply for military service. Enlistment models have generally dealt with explaining the variation observed in a ratio of applicants and/or enlistment contracts to an age specific youth population, generally 17 to 21 year olds, per time period (Cotterman, 1986; DeVany and Saving, 1982; Goldberg, 1988; Saving and Stone, 1983; Saving, Stone, Looper, and Taylor, 1985; Stone, Saving, Turner, and Looper, 1992). The analysis models which are presented in the following sections model the decisions of youth to apply for military service as a function of demographic and propensity measures. These demographic and propensity measures are derived from responses provided to the YATS II surveys performed between 1984 to 1989.

#### **Matching YATS and MEPS Data**

To determine whether the individual YATS respondent applied for military service, data from the YATS were matched with MEPS applicant data by the Defense Manpower Data Center (DMDC) using the Social Security Administration Number (SSAN) provided by respondents to the survey. YATS respondents are matched with MEPS historical files both prior to and after the date on which the YATS was administered to insure finding the earliest possible MEPS applicant record match. For each YATS respondent whose SSAN matched with a MEPS applicant record, the amount of lapsed time between the date the YATS survey was administered and the date the individual applied at the MEPS was calculated, DIFFDT. Early analysis of the matched YATS and MEPS data indicated that some values for DIFFDT were negative.

DIFFDT was calculated using the DATE OF EARLIEST APPLICATION from the MEPS applicant records (date the individual applied at the MEPS) and the interview date *INTDATE* from the YATS. A distribution of the calculated DIFFDT values for YATS data from 1984 to 1989 is presented in Table 1. Each year of the YATS was matched against MEPS applicant records from fiscal year 1980 (FY80) to FY91. As the distributions indicate, many of the YATS respondents who matched the MEPS had an application date prior to responding

to the YATS (negative values for DIFFDT). This implies that the respondent had begun the MEPS process (based on the application date from the MEPS file) prior to the date of the YATS interview.

**Table 1. Distribution of DIFFDT Using MEPS Date**

<b>Year/DIFFDT</b>	<b>1984</b>	<b>1985</b>	<b>1986</b>	<b>1987</b>	<b>1988</b>	<b>1989</b>
<b>less than -23 months</b>	57 (7.0)	121 (12.9)	159 (18.6)	190 (20.7)	234 (26.1)	286 (27.2)
<b>-23 months to -12 months</b>	91 (11.3)	77 (8.2)	66 (7.7)	76 (8.3)	87 (9.7)	89 (8.5)
<b>-11 months to -1 months</b>	93 (11.5)	104 (11.1)	89 (10.4)	97 (10.6)	87 (9.7)	130 (12.3)
<b>Total with less than 0 months</b>	<b>241 (29.8)</b>	<b>302 (32.2)</b>	<b>314 (36.6)</b>	<b>363 (39.6)</b>	<b>408 (45.4)</b>	<b>505 (48.0)</b>
<b>0 months</b>	29 (3.6)	26 (2.8)	22 (2.6)	20 (2.2)	30 (3.3)	42 (4.0)
<b>1 month to 6 months</b>	109 (13.5)	119 (12.7)	122 (14.2)	115 (12.6)	121 (13.5)	178 (16.9)
<b>7 months to 18 months</b>	183 (22.6)	206 (22.0)	183 (21.4)	216 (23.6)	196 (21.8)	238 (22.6)
<b>19 months to 30 months</b>	105 (13.0)	116 (12.4)	105 (12.3)	133 (14.5)	101 (11.2)	89 (8.5)
<b>31 months to 42 months</b>	55 (6.8)	80 (8.5)	70 (8.2)	52 (5.7)	42 (4.7)	N/A
<b>over 42 months</b>	87 (10.8)	88 (9.4)	41 (4.8)	17 (1.9)	N/A	N/A
<b>Total</b>	<b>809</b>	<b>937</b>	<b>857</b>	<b>916</b>	<b>898</b>	<b>1,052</b>

Not all respondents who have a MEPS application date actually enlist in the military. Further analysis of the MEPS data indicates that a significant proportion of the MEPS applicants never agree to enlist. Reasons for failure to pursue the MEPS contact could be numerous: failed physical, low ASVAB scores, morally unqualified, lost interest, etc. To determine whether the individual had made a formal commitment to the military before responding to the YATS survey, three dates from the MEPS records were used: DELAYED ENTRY PROGRAM

(DEP) IN DATE (date the individual entered the DEP), DATE OF CONTRACT (date on which the individual signed a contract to enter the military), and DATE OF ENTRY (date on which the individual was to enter the military). SIGNDT was calculated using the earliest of these three dates and the YATS interview date. The results are shown in Table 2.

**Table 2. Distribution of SIGNDT**

<b>Year/SIGNDT</b>	<b>1984</b>	<b>1985</b>	<b>1986</b>	<b>1987</b>	<b>1988</b>	<b>1989</b>
<b>less than -23 months</b>	5 (0.6)	10 (1.1)	20 (2.3)	18 (2.0)	34 (3.8)	37 (3.5)
<b>-23 months to -12 months</b>	9 (1.1)	6 (0.6)	14 (1.6)	6 (0.7)	15 (1.7)	14 (1.3)
<b>-11 months to -1 months</b>	8 (1.0)	10 (1.1)	5 (0.6)	6 (0.7)	5 (0.6)	5 (0.5)
<b>Total with less than 0 months</b>	<b>22</b> <b>(2.7)</b>	<b>26</b> <b>(2.8)</b>	<b>39</b> <b>(4.6)</b>	<b>30</b> <b>(3.3)</b>	<b>54</b> <b>(6.0)</b>	<b>56</b> <b>(5.3)</b>
<b>0 months</b>	8 (1.0)	10 (1.1)	9 (1.1)	7 (0.8)	6 (0.7)	19 (1.8)
<b>1 month to 6 months</b>	76 (9.4)	72 (7.7)	71 (8.3)	75 (8.2)	58 (6.5)	108 (10.3)
<b>7 months to 18 months</b>	109 (13.5)	137 (14.6)	116 (13.5)	141 (15.4)	128 (14.3)	160 (15.2)
<b>19 months to 30 months</b>	96 (11.9)	90 (9.6)	65 (7.6)	110 (12.0)	84 (9.4)	73 (6.9)
<b>31 months to 42 months</b>	56 (6.9)	49 (5.2)	60 (7.0)	48 (5.2)	34 (3.8)	N/A
<b>over 42 months</b>	73 (9.0)	71 (7.6)	28 (3.3)	14 (1.5)	N/A	N/A
<b>MEPS Date only</b>	<b>369</b> <b>(45.6)</b>	<b>482</b> <b>(51.4)</b>	<b>469</b> <b>(54.7)</b>	<b>491</b> <b>(53.6)</b>	<b>534</b> <b>(59.5)</b>	<b>636</b> <b>(60.5)</b>
<b>Total</b>	<b>809</b>	<b>937</b>	<b>857</b>	<b>916</b>	<b>898</b>	<b>1,052</b>

As Table 2 indicates, the distribution of SIGNDT is significantly different from the distribution of DIFFDT presented in Table 1. Few respondents exhibited a negative SIGNDT, most respondents with a negative DIFFDT only possessed a MEPS date. It can be implied from

the SIGNDT distribution that a YATS respondent with a negative value has already made a formal commitment to the military before responding to the YATS. A record with only a MEPS date suggests that the individual's application has not yet resulted in a formal contract at this point.

Individuals with a negative DIFFDT, and especially those with a negative SIGNDT, present a potential problem when developing a model to predict the propensity of youth to apply to the military. Individuals with a negative DIFFDT had already applied prior to responding to the YATS. Having already begun the MEPS process and made the decision to apply to the military, these individuals' responses to the YATS may be biased relative to other respondents who have not made a decision to apply for military service. The same holds even more so for respondents who have already made a formal commitment to the military before taking the YATS. For the purpose of this study, individuals with a negative DIFFDT and individuals with a negative SIGNDT will be excluded from the analysis. It is assumed that the inclusion of those individuals who had already made a decision to apply to the military will bias the results of the estimation of the propensity model.

### **Sample Population**

Respondents to the YATS are prompted for their SSANs at the end of the survey. Providing their SSAN is voluntary for respondents to the YATS. Matching YATS respondents with the MEPS historical files requires that respondents have provided an SSAN. Without providing a SSAN, it is not possible to determine whether a respondent did or did not apply for military service. For this study, respondents without an SSAN are excluded from the analysis sample.

For the surveys performed between 1984 and 1989, this excludes approximately 36.14% of the total respondents, though the percentage does tend to vary by YATS survey as indicated in Table 3. Though the percentage of respondents providing a SSAN tends to be relatively stable over the 1984 to 1989 time period, it does fluctuate from a low of 59.89% in 1986 to a high of 69.80% in 1989. Thus, over the six year sample, the percentage of SSAN providers averages approximately 63.69%.

**Table 3. SSANs Provided by Year of YATS**

<b>Year</b>	<b>Total Respondents</b>	<b>Percent with SSANs</b>
<b>1984</b>	7,940	66.52
<b>1985</b>	9,959	65.03
<b>1986</b>	10,743	59.89
<b>1987</b>	11,271	60.62
<b>1988</b>	10,985	61.53
<b>1989</b>	11,575	69.81
<b>Weighted Average</b>	10,412	63.69

Females became an increasingly larger percentage of the YATS sample population from 1984 to 1989 as indicated in Table 4. This table also shows that females appear to be only slightly less likely to provide their SSANs than male respondents. This increased interest by DoD in female attitudes toward military service was engendered by their growing roles in the military which fostered the need for information on the female population concerning their

**Table 4. SSANs Provided by Females by Year of YATS**

<b>Year</b>	<b>Number of Females</b>	<b>Number of Males</b>	<b>Percent Female</b>	<b>Percent Female with SSAN</b>	<b>Percent Male with SSAN</b>
<b>1984</b>	1,503	6,437	18.93	64.40	65.90
<b>1985</b>	3,301	6,658	33.14	62.90	66.07
<b>1986</b>	4,293	6,450	39.96	59.38	60.23
<b>1987</b>	4,525	6,745	40.15	58.28	62.19
<b>1988</b>	4,369	6,616	39.77	60.20	62.41
<b>1989</b>	4,596	6,979	39.71	67.12	71.57

propensity for military service. Female respondents from the YATS will be included in the estimation of the application models of the PPMM.

## IV. REPLICATION AND EXTENSION OF EARLIER INTENTION ANALYSIS

The Orvis studies used the match between the YATS survey respondents and the MEPS applicant records as the basis for their analysis. One focus of the Orvis research was the relationship between the stated intention of YATS respondents to enlist in the military and their actual enlistment behavior. Orvis used two questions from the YATS (Q438 and Q503, see Appendix C) to construct an intention to enlist measure for each YATS respondent. *Question 438* asks the respondent what his plans for the next few years are in the context of going to school, working, or joining the military (i.e., occupational choices). If the respondent mentions joining the military, he is considered an *unaided mention* for military service. *Question 503* asks the respondent how likely it is that he will be serving in the military in the next few years. The possible responses are "definitely," "probably," "probably not," or "definitely not"; in the few instances when the respondent cannot decide, he may indicate "don't know" as a response. Orvis used these two questions from the survey to divide respondents into four propensity groups based on their stated intention to serve in the military:

*Group 1. Unaided mention, definite intention*

*Group 2. Unaided mention, probable intention*

*Group 3. No unaided mention, definite or probable intention, and*

*Group 4. No unaided mention, probably not or definitely not intention.*

Orvis included the "don't know" respondents in Group 4, the negative propensity group. In the present sample, this group comprises less than 0.2% of total YATS respondents.

### Variable Description

To replicate the work of Orvis with respondents to later YATS requires some modification in the specification of the equation since some of the variables originally used by Orvis are no longer collected in the YATS surveys. Where possible, proxies are used. The application equation is estimated using a logit function, since the dependent variable has a binary zero/one value. A value of one implies that the respondent applied for military service (matched

the MEPS applicant records), and zero otherwise. Table 5 details explanatory variables used in the replication of the Orvis work. All of the variables in Table 5 are binary with the exception of Q403 (age), Q700 (grades), and Q713F (father's education). A more detailed explanation of the variables and how they were constructed from the YATS questions is presented in Appendix C.

### **Estimation of the Application Equation**

Table 6 presents the results of the replication of the Orvis work. The equation was estimated using a logit function over the 1984 to 1987 YATS respondents. Throughout this report all logit coefficients have been converted to probabilities and these probabilities are presented in the tables of estimation results. Only 16 to 21 year-old male respondents are included in the equation. As indicated in Table 6, the three intention variables are directly related to the propensity to apply and statistically significant at the 99% level, which is consistent with the Orvis findings. Several other variables are statistically significant at the 99% level:

- (1) Q403 (Age) - age is inversely related to the propensity to apply, i.e., as the individual gets older, the propensity to apply declines approximately 0.022 points per year of age,
- (2) DQ709 (Physics) - respondents who have taken physics exhibit approximately 0.040 points lower propensity to apply than those who have not,
- (3) Q700 (Grades) - grades are inversely related to the propensity to apply, e.g., the propensity to apply decreases approximately 0.018 points each 0.5 average grade point drop,
- (4) DFJOB - respondents who perceive no difficulty finding a job exhibit approximately 0.032 points lower propensity to apply than those who do perceive difficulty, and
- (5) DRECR - respondents who have contacted a recruiter exhibit approximately 0.144 points higher propensity to apply than those who have not.

**Table 5. Variable Definitions**

<b>Variable Name</b>	<b>Description</b>
<b>PMIL1</b>	Unaided mention, definite intention
<b>PMIL2</b>	Unaided mention, probable intention
<b>PMIL3</b>	No unaided mention, definite or probable intention
<b>Q403</b>	Age: Values 16 to 21
<b>DSOUTH</b>	Census district South
<b>DNORTH</b>	Census district North Central
<b>DWEST</b>	Census district West
<b>DBLACK</b>	Black
<b>DOTHER</b>	Non-Caucasian and non-black
<b>DDIPLOMA</b>	High school diploma
<b>DISCHOOL</b>	Still attending school
<b>DQ702</b>	Taken elementary algebra
<b>DQ703</b>	Taken plane geometry
<b>DQ705</b>	Taken computer science
<b>DQ706</b>	Taken intermediate algebra
<b>DQ707</b>	Taken trigonometry
<b>DQ708</b>	Taken calculus
<b>DQ709</b>	Taken physics
<b>Q700</b>	Grades in school: Value 1 to 7
<b>Q713F</b>	Father's education
<b>DCWORK</b>	Currently working
<b>DLWORK</b>	Currently not employed and looking for work
<b>DFJOB</b>	No difficulty finding a job
<b>DRECR</b>	Have contacted a recruiter

**Table 6. Results of the Replication of Earlier Propensity Models**

	Number of Observations	10,906		
	Log Likelihood	-5095.42		
	Chi-Square	1107.70		
<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.2131
PMIL1	0.2456	12.781	0.000	0.0349
PMIL2	0.2334	14.616	0.000	0.0521
PMIL3	0.1169	10.974	0.000	0.1691
Q403	-0.0221	-5.711	0.000	18.0932
DSOUTH	0.0164	1.350	0.177	0.3545
DNORTHC	0.0286	2.247	0.025	0.2657
DWEST	0.0226	1.604	0.109	0.1815
DBLACK	0.0172	1.335	0.182	0.1093
DOTHER	-0.0227	-1.481	0.138	0.0899
DDIPLOMA	0.0027	0.242	0.808	0.4681
DISCHOOL	-0.0158	-1.498	0.134	0.6585
DQ702	0.0234	1.911	0.056	0.8165
DQ703	-0.0104	-0.982	0.326	0.6241
DQ705	-0.0044	-0.447	0.655	0.3058
DQ706	0.0010	0.093	0.926	0.5606
DQ707	0.0045	0.362	0.717	0.2878
DQ708	-0.0341	-1.989	0.047	0.1151
DQ709	-0.0405	-3.477	0.001	0.2384
Q700	0.0181	5.020	0.000	3.4651
Q713F	-0.0005	-0.276	0.782	12.8226
DCWORK	-0.0022	-0.168	0.866	0.6447
DLWORK	0.0149	1.029	0.303	0.1993
DFJOB	-0.0322	-3.053	0.002	0.2312
DRECR	0.1440	16.300	0.000	0.4651
CONSTANT	-0.0118	-0.156	0.876	1.0000

Other explanatory variables in the estimation equation are significant at the 90% level of confidence or above. These variables included:

- (1) DNORTHC - respondents who are from the North Central census district exhibit approximately 0.029 points higher propensity to apply than those who are not,
- (2) DQ702 (Elementary Algebra) - respondents who have taken elementary algebra exhibit approximately 0.023 points higher propensity to apply than those who have not, and

- (3) DQ708 (Calculus) - respondents who have taken calculus exhibit approximately 0.034 points lower propensity to apply than those who have not.

These results are consistent with results previously published by Orvis though the propensity gains and losses exhibited by the coefficients are different in magnitude.

### **Preliminary Extension of the Application Equation**

In a preliminary extension of the Orvis work, the number of intention groups was expanded. Initially, intention groups were divided into four groups based on *Questions 428 and 503*. Using the same questions as the base, group three from before was split into two groups, and two additional groups were defined. The expanded intention groups are defined as:

- Group 1. Unaided mention, definite intention,*
- Group 2. Unaided mention, probable intention,*
- Group 3. No unaided mention, definite intention,*
- Group 4. No unaided mention, probable intention,*
- Group 5. No unaided mention, probably not intention, and*
- Group 6. No unaided mention, definitely not intention.*

These additional categories of intention are represented as explanatory variables PMIL4 and PMIL5 in the estimation equation. Group six is included in the intercept for the estimation. A more detailed description of how the new intention variables were constructed is provided in Appendix C.

The results of adding the new explanatory intention variables are presented in Table 7. Both of the new intention categories are statistically significant at the 99% level of confidence. The coefficients of the PMIL1, PMIL2, and PMIL3 groups exhibit changes in size, i.e., compared to the negative intention group PMIL1 indicates that respondents who have an "unaided mention" and "definite" intention exhibit approximately 0.268 (compared to 0.246 in Table 6) points higher propensity to apply, approximately 0.255 (compared to 0.233 in Table 6) points higher propensity to apply for PMIL2, and approximately 0.211 (compared to 0.117 in Table 6) points higher propensity to apply for PMIL3.

**Table 7. Results Including New Intention Measure**

Number of Observations	10,906
Log Likelihood	-5082.60
Chi-Square	1133.34

<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.2131
PMIL1	0.2680	13.407	0.000	0.0349
PMIL2	0.2554	15.125	0.000	0.0521
PMIL3	0.2112	7.963	0.000	0.0181
PMIL4	0.1291	10.454	0.000	0.1510
PMIL5	0.0429	4.065	0.000	0.3239
Q403	-0.0212	-5.467	0.000	18.0932
DSOUTH	0.0155	1.275	0.202	0.3545
DNORTH	0.0293	2.303	0.021	0.2657
DWEST	0.0222	1.568	0.117	0.1815
DBLACK	0.0184	1.424	0.154	0.1093
DOTHER	-0.0237	-1.545	0.122	0.0899
DDIPLOMA	0.0044	0.387	0.698	0.4681
DISCHOOL	-0.0169	-1.596	0.111	0.6585
DQ702	0.0238	1.942	0.052	0.8165
DQ703	-0.0104	-0.981	0.326	0.6241
DQ705	-0.0036	-0.370	0.711	0.3058
DQ706	0.0007	0.066	0.948	0.5606
DQ707	0.0053	0.430	0.667	0.2878
DQ708	-0.0323	-1.879	0.060	0.1151
DQ709	-0.0423	-3.628	0.000	0.2384
Q700	0.0184	5.064	0.000	3.4651
Q713F	-0.0003	-0.176	0.860	12.8226
DCWORK	-0.0026	-0.202	0.840	0.6447
DLWORK	0.0139	0.956	0.339	0.1993
DFJOB	-0.0316	-2.988	0.003	0.2312
DRECR	0.1418	16.026	0.000	0.4651
CONSTANT	-0.0503	-0.656	0.511	1.0000

### **Including Males and Females in the Application Equation**

As indicated earlier in Section III, the YATS sample has continually been expanded to include a significant number of females. Table 8 shows the increasing proportion of the sample of 16 to 21 year-old YATS respondents providing a SSAN that are female. To account for this increased sampling of the female population and the increased roles of females in the military, females respondents were added to the estimation sample.

**Table 8. Males and Females (16 to 21) in the YATS**

<b>Year</b>	<b>Number of Respondents with SSAN</b>	<b>16 to 21 Percent Male</b>	<b>16 to 21 Percent Female</b>
<b>1984</b>	4,309	77.23	22.46
<b>1985</b>	5,653	63.26	36.74
<b>1986</b>	5,034	62.85	37.15
<b>1987</b>	5,485	63.57	36.43
<b>1988</b>	5,337	63.42	36.57
<b>1989</b>	6,475	63.88	36.12

Table 9 provides the results of estimating the equation in Table 7 using both male and female respondents. The binary variable DMALE (1 if male, 0 if female) was added to the equation to account for the average difference between male and female propensity to apply for military service. The coefficient for DMALE which is statistically significant at the 99% level implies that males exhibit approximately 0.133 points higher propensity to apply than females. All the intention variables are statistically significant at the 99% level and experience some change in the size of their coefficients. The addition of females to the sample affects several of the variables compared to the Table 7 results for males only:

- (1) PMIL1 - unaided definite intention respondents exhibit approximately 0.210 points higher propensity to apply than the negative (definitely not) intention group (compared to 0.268 points in Table 7, a 21.6% decrease),
- (2) PMIL2 - unaided probable intention respondents exhibit approximately 0.191 points higher propensity to apply than the negative (definitely not) intention group (compared to 0.255 points in Table 7, a 25.1% decrease),
- (3) PMIL3 - positive (definite) intention respondents exhibit approximately 0.143 points higher propensity to apply than the negative (definitely not) intention group (compared to 0.211 points in Table 7, a 32.2% decrease),

**Table 9. Results for Male and Female Equation**

Number of Observations	10,484
Log Likelihood	-5442.54
Chi-Square	2119.92

<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.1343
DMALE	0.1334	19.184	0.000	0.4999
PMIL1	0.2096	14.669	0.000	0.0222
PMIL2	0.1910	15.751	0.000	0.0329
PMIL3	0.1425	7.574	0.000	0.0132
PMIL4	0.1042	12.014	0.000	0.1067
PMIL5	0.0372	5.136	0.000	0.2714
Q403	-0.0114	-4.340	0.000	18.1396
DSOUTH	0.0118	1.393	0.164	0.3613
DNORTH	0.0174	1.958	0.050	0.2622
DWEST	0.0140	1.423	0.155	0.1813
DBLACK	0.0164	1.879	0.060	0.1139
DOTHER	-0.0042	-0.402	0.687	0.0949
DDIPLOMA	0.0104	1.347	0.178	0.4902
DISCHOOL	-0.0100	-1.379	0.168	0.6451
DQ702	0.0106	1.251	0.211	0.8270
DQ703	-0.0068	-0.938	0.348	0.6162
DQ705	-0.0007	-0.109	0.913	0.2888
DQ706	0.0141	1.972	0.049	0.5585
DQ707	0.0012	0.138	0.890	0.2685
DQ708	-0.0405	-3.266	0.001	0.1065
DQ709	-0.0280	-3.398	0.001	0.2041
Q700	0.0110	4.346	0.000	3.2411
Q713F	-0.0014	-1.278	0.201	12.7642
DCWORK	0.0054	0.624	0.533	0.6149
DLWORK	0.0011	0.108	0.914	0.2079
DFJOB	-0.0223	-2.980	0.003	0.2109
DRECR	0.1044	17.134	0.000	0.3663
CONSTANT	-0.2254	-4.307	0.000	1.0000

\* Separate models for males and females will be developed in Section V.

- (4) PMIL4 - positive (probable) intention respondents exhibit approximately 0.104 points higher propensity to apply than the negative (definitely not) intention group (compared to 0.129 points in Table 7, a 19.4% decrease),
- (5) PMIL5 - negative (probably not) intention respondents exhibit approximately 0.037 points higher propensity to apply than the

negative (definitely not) intention group (compared to 0.043 points in Table 7, a 14.0% decrease),

- (6) Q403 - the coefficient for age decreased 46.0% from Table 7, indicating that the propensity to apply decreases 0.011 points for each additional year of age when females are included in the estimation,
- (7) DBLACK - became statistically significant at the 90% level of confidence, black respondents exhibit approximately 0.016 points higher propensity to apply than caucasians,
- (8) DQ702 - became statistically insignificant, i.e., respondents who had taken elementary algebra are not statistically different from those who had not in their propensity to apply,
- (9) Q700 - the propensity to apply decreases approximately 0.011 points each 0.5 average grade point drop (compared to 0.018 points in Table 7, a 38.9% decrease),
- (10) DFJOB - respondents who have no perceived difficulty of finding a job exhibit approximately 0.022 points lower propensity to apply than those who do (compared to 0.032 points in Table 7, a 31.3% decrease), and
- (13) DRECR - respondents who had contacted a recruiter exhibit approximately 0.104 points higher propensity to apply than those who had not (compared to 0.142 points in Table 7, a 26.8% decrease).

## V. PPMM ESTIMATION

The estimation equations of the Population Propensity Measurement Model (PPMM) include YATS respondents from the 1984 to 1987 surveys. These are the same survey years as used in the replication and expansion of the Orvis estimations presented in Section IV. As in the replication estimations, only YATS respondents with a SSAN are included in the estimation sample. However, for the PPMM estimation, respondents with a negative DIFFDT or SIGNDT (from Section III) are excluded from the analysis sample. Each equation has variables comprised of attitudinal and demographic factors. Equations are estimated for four youth groups:

*Males, 16 to 18 years of age,*

*Males, 19 to 21 years of age,*

*Females, 16 to 18 years of age, and*

*Females, 19 to 21 years of age.*

The dependent variable used in the analyses is binary, and equal to zero if the respondent did not apply for military service, or one if the respondent applied for military service. A respondent is considered to have applied for military service if his SSAN matches with a MEPS application record. The estimated models provide the basis for projecting the propensity to enlist for future youth populations sampled by YATS. If the YATS sample is assumed to be representative of youth propensities to apply for military service, the population weights calculated for YATS samples can then be used to extrapolate to the youth population, in general, and, thus, project application rates based on future YATS survey samples.

Since the dependent variable has a binary zero/one value, the application equations are estimated using a logit function. The logit function is mathematically constrained to always predict probability values between zero and one. In addition, logit possesses desirable large sample estimation properties which are not maintained by ordinary least squares when estimating an equation with a binary dependent variable (Theil, 1971).

### **Intensity of Intention on the Propensity to Apply**

The initial intention measures constructed by Orvis, have been further expanded to include additional questions. Active-duty Service specific intention questions (*Q510* - Army, *Q511* - Air Force, *Q512* - Marine Corps, and *Q513* - Navy) and Reserve and Guard component specific intention questions (*Q505* - National Guard, *Q507* - Reserves, and *Q509* - Coast Guard) are combined with the initial questions *Q438* and *Q503* to construct new intention variables. Using the responses to these six questions allows a better perception of the respondents overall intensity or sincerity to apply for military service. Responses to *Questions 510* through *513* are the same responses as to *Q503*: "definitely," "probably," "probably not," and "definitely not." The new intention variables are represented in Table 10. A more detailed description of how the new intention variables were created is provided in Appendix C. Questions from the YATS interviews used to construct these variables are provided in Appendix D.

**Table 10. New Intention Variables**

<b>Variable Name</b>	<b>Description</b>
<b>PMIL41</b>	Unaided mention, definite or probable intention, and definite or probable intention for at least one of the four active-duty Services or Reserve or Guard components
<b>PMIL42</b>	No unaided mention, definite or probable intention, and definite or probable intention for at least one of the four active-duty Services or Reserve or Guard components
<b>PMIL43</b>	No unaided mention, probable or probably not intention, and probable, probably not, or definitely not intention for at least one of the four active-duty Services or Reserve or Guard components
<b>PMIL44</b>	No unaided mention, definitely not intention, and probably not or definitely not intention for at least one of the four active-duty Services or Reserve or Guard components

### **Potential Biases in the Application Data**

An additional factor which appears to affect the respondents' propensity to apply for military service is whether the respondent has contacted a recruiter or taken the ASVAB. The location of where the ASVAB was taken, at a MEPS, at a Mobile Examining Team Sight (METS), or in high school, also appears to affect the propensity to apply. To account for the

impact of contacting a recruiter and/or taking the ASVAB upon enlistment propensity, five new variables are added to the analysis (Table 11). These variables are constructed from *Questions* 628, 645, and 647. The percentage comprising each of these groups increased slightly over the 1984 to 1989 time period as exhibited in Table 12, with the exception of respondents taking the ASVAB at a MEPS or METS. Table 13 shows the percentage of respondents in each of these five groups (Table 11) that eventually apply for military service. Respondents in these five groups consistently display higher application rates than respondents not included in any of the five groups. For a more detailed description of how the new recruiter and ASVAB variables were created, see Appendix C. Questions from the YATS interviews used to construct these variables are provided in Appendix D.

**Table 11. New Variables for ASVAB-Taking and Recruiter Contact**

Variable Name	Description
DRECR2	Contacted a recruiter but has not taken the ASVAB
DASVAB	Has not contacted a recruiter but has taken the ASVAB
DAPPEP	Contacted a recruiter and taken the ASVAB at the MEPS
DAPMET	Contacted a recruiter and taken the ASVAB at a METS
DAPPHS	Contacted a recruiter and taken the ASVAB in high school

### **Schooling and Non-Schooling Groups of Respondents**

Orvis used two variables to account for the impact of being in school and having a high school diploma on the propensity to apply. In the equation presented in Section IV, the variables DISCHOOL and DDIPLOMA are used to represent these two effects. Approximately 49.0% of the sample have high school diplomas, while 64.5% of the sample are still in school. As expected, the largest proportion of the sample which is still in school is in the 16 to 18 year old age group. The largest proportion of the sample which has a high school diploma is in the 19 to 21 year old age group.

**Table 12. Percent Contacted Recruiter and Taken ASVAB**

Year	Percent Contacted Recruiter (Only)	Percent Taken ASVAB (Only)	Percent Contacted Recruiter and Taken ASVAB at a MEPS	Percent Contacted Recruiter and Taken ASVAB at a METS	Percent Contacted Recruiter and Taken ASVAB in High School
1984	27.77	6.07	2.29	0.56	7.87
1985	25.23	7.63	1.59	0.41	8.29
1986	25.50	7.52	1.40	0.36	8.32
1987	28.43	7.27	1.47	0.35	10.53
1988	27.90	8.83	1.35	0.66	10.09
1989	29.64	8.59	1.42	0.46	10.86

**Table 13. Percentage Applied by Recruiter and ASVAB Groups**

Year	Percent Applied & DRECR2	Percent Applied & DASVAB	Percent Applied & DAPPMET	Percent Applied & DAPPMET	Percent Applied & DAPPHS	Percent Applied & None
1984	17.6	16.6	27.9	21.4	19.1	9.7
1985	17.3	10.0	15.6	15.0	18.2	8.0
1986	16.5	9.5	25.4	18.8	16.8	7.2
1987	13.3	9.1	14.1	29.4	16.0	7.6
1988	13.0	7.8	12.6	29.0	15.5	6.0
1989	13.9	5.4	17.5	11.5	11.0	5.0

Potential differences might be expected in the propensity to apply between 19 to 21 year-old respondents and 16 to 18 year-old respondents who are still in high school. In addition, differences might also be expected in the propensity to apply between 16 to 18 year-old respondents and 19 to 21 year-old respondents who have a high school diploma. These potential differences are analyzed using *Questions 404* and *406* from the YATS. *Q404* asks for the highest grade completed in school. If the response to *Q404* is 12 or greater (completed high

school or greater), then the respondent is asked *Q406*, which asks for the type of degree received, if any, e.g., high school diploma, GED, ABE, etc.

*Q403* was used to construct age specific binary variables for the estimation of PPMM to determine if there are differences in the propensity to apply by age. To account for these potential differences, eight new variables were formed based on age and whether the respondent completed high school. The new variables for schooling are presented in Table 14. The variables D16I, D17I, D18I, D19O, and D20O are binary variables which take the value 1 for the appropriate age (suffix I for in high school and suffix O for graduated high school), 0 otherwise. For a more detailed description of these new variables and the YATS questions from which they were created, see Appendix C. Questions from the YATS interviews used to construct these variables are provided in Appendix D.

**Table 14. New Variables for Schooling**

Variable Name	Description
D16I	Age 16 and still in high school
D17I	Age 17 and still in high school
D18I	Age 18 and still in high school
D19O	Age 19 and have completed high school
D20O	Age 20 and have completed high school
D1618G	Age 16 to 18 and have completed high school
D1618NG	Age 16 to 18 and not in high school but have not received a diploma of any type
D1921I	Age 19 to 21 and have not completed high school and have not received a diploma of any type

### **Other Explanatory Variables from YATS**

*Question 436* was used to construct the variable DFJOB. This variable measures the respondent's perceived difficulty of finding a job in the community and is used as a proxy for local unemployment rates. DFJOB is constructed in the same manner as in the Orvis replication (Section IV and Appendix C). Several other factors which affect the decision to apply for

military service can be created from the responses to YATS. These variables are presented in Table 15. DMILCOL is a binary variable that takes a value of one if the respondent said he was "definitely" going to college in *Q514* and then said he was "planning to serve in the military" at *Q517*. This binary variable is attempting to account for a group of respondents who plan to attend college after serving in the military.

**Table 15. Other Explanatory Variable Definitions**

Variable Name	Description
DMILCOL	Definite intention for college and responded military
DQ709P	Taken or plan to take physics
Q700x2	Value of <i>Q700</i> * Value of <i>Q700</i>
Q713Fx2	Value of <i>Q713F</i> * Value of <i>Q713F</i>

As in the Orvis replication and extension, the respondent's grades (*Q700*) and father's education (*Q713F*) are included in the estimation equations. However, preliminary analysis of the data using neural networks suggested a nonlinear relationship existed between grades and applying for military service and also between father's education and applying for military service. Therefore, grades and father's education have been squared (*Q700x2* and *Q713Fx2*) and added to the estimation equations to account for this nonlinear relationship.

Whether the respondent had taken the courses in questions *Q702* through *Q708* were found not to be statistically significantly related to the decision to apply for military service, and these questions were therefore dropped from the estimation equation. Having completed a course in physics (*Q709*), was found to impact the decision to apply. Improved response from this course is found when the binary variable for the courses is reconstructed to include respondents who "planned" to take the course with respondents who had taken the course. The reconstructed variable is *DQ709P* and it has been added to the estimation equations. For a more detailed description of the YATS questions used to create these variables, refer to Appendix C. Questions from the YATS interviews used to construct these variables are provided in Appendix D.

## The Importance of the Population Distribution

The equations presented in Section IV were comprised of demographic and intention information concerning YATS respondents. The average propensity to apply or enlist in the military has not changed significantly according to the measures constructed for the YATS based on responses to intention questions (*CPYATS82*, *RSVNG84*, *Q503*, and *Q522*). As Table 16 indicates, the two composite propensity measures, *CPYATS82* and *RSVNG84* varied less than 0.7% and 0.6%, respectively over the 1984 to 1989 time period. Two of the intention questions, *Q503* and *Q522*, exhibited variations of 0.5% and 1.2%, respectively. Conversely, the application rate did vary over the 1984 to 1989 time period, though the last two years may still experience increases as the 16 to 18 year-olds become eligible for service or look to the military as a viable alternative for employment and/or training, especially 16 to 18 year-old respondents in 1988 and 1989. The application rate (determined from YATS respondents providing valid SSANs) increased every year from 1984 to 1989 (Table 17).

Table 16. Variation in Intention Over Time

Year	CPYATS82	RSVNG84	Q503	Q522
1984	3.1417	3.2828	3.2386	2.8930
1985	3.1542	3.2853	3.2809	2.8428
1986	3.1532	3.3399	3.2541	2.7896
1987	3.1398	3.3193	3.2611	2.8187
1988	3.2066	3.3597	3.2954	2.6814
1989	3.1213	3.2624	3.2237	2.9287

If the application rate does not vary sufficiently over the estimation time period, 1984 to 1987, then the estimated equation for PPMM may not provide movements in future forecasts, i.e., the equation would produce a constant application rate. To engender fluctuation in the yearly forecasts, the YATS explanatory variables used to estimate PPMM must reflect fluctuations which occur annually and are related to the propensity to enlist. To induce more movement in the in-sample and out-of-sample predictions, a new variable is added to the list of

**Table 17. Variation in the Application Rate Over Time**

Year	Actual Application Rate	Change in Actual Application Rate	Percent Change in Actual Application Rate
1984	0.1355	-----	-.--
1985	0.1149	-0.0206	-15.18
1986	0.1083	-0.0066	-5.78
1987	0.1037	-0.0046	-4.28
1988	0.0931	-0.0106	-10.23
1989	0.0850	-0.0080	-8.63

explanatory variables to account for the general distribution of population, the population of 17 to 21 years male and females at the county FIPS by year (SEMA1721). SEMA1721 is expected to vary inversely with the application rate. The inverse relationship implies that low population areas tend to exhibit relatively higher application rates than high population areas. Data from the YATS presented in Table 18 tend to support this inverse relationship.

**Table 18. Population Variation Compared with Application Rates**

Year	Pop < 2,500	Pop > 2,500 Pop < 10,000	Pop > 10,000 Pop < 25,000	Pop > 25,000 Pop < 100,000	Pop > 100,000
1984	0.1312	0.1449	0.1150	0.1177	0.1034
1985	0.1286	0.1188	0.1031	0.0812	0.0885
1986	0.1000	0.1198	0.1031	0.0884	0.0840
1987	0.1097	0.0982	0.0772	0.0858	0.0556
1988	0.0872	0.1002	0.0790	0.0803	0.0405
1989	0.0894	0.0833	0.0740	0.0635	0.0679
Average	0.1053	0.1088	0.0909	0.0859	0.0761

## **Estimation Results for PPMM**

The PPMM estimation results are presented for each of the male and female age groups. Results for males 16 to 18 years of age and 19 to 21 years of age are presented in Tables 19 and 20, respectively. Coefficients vary widely between the two male age groups. The application rate is approximately 3 times higher for younger males (16 to 18) than it is for older males (19 to 21). The PMIL intention variables are all significant at the 99% level of confidence in the male equations. The size of the coefficients of the PMIL intention variables all decrease with level of stated intention. This is true for both the younger and older male equations. Perceived difficulty finding a job (DFJOB), was statistically significant at the 90% level of confidence in both male equations. Grades (Q700) appear to be important for younger males, but statistically insignificant in the older male equation. DMILCOL appears to be important in determining enlistment behavior among younger males, but unimportant for the behavior of older males. The level of the father's education also appears to be more important in determining younger males' enlistment behavior than older males' behavior. Binary variables accounting for contacting a recruiter and/or taking the ASVAB are statistically significant at the 99% level in both equations, with the exception of DAPPMET (significant at the 90% level of confidence). Comparison of large differences in the statistically significant coefficients (95% or above level of confidence for both coefficients) of the equations for males, 16 to 18 years of age versus males, 19 to 21 years of age (Tables 19 and 20) are:

- (1) DFJOB - male respondents, 16 to 18 years of age perceiving no difficulty finding a job exhibit approximately 0.028 points lower propensity to apply than respondents perceiving difficulty finding a job, 34.4% higher than positive intention male respondents, 19 to 21 years of age (-0.018),
- (2) PMIL41 - male respondents, 16 to 18 years of age exhibit approximately 0.280 points higher propensity to apply than the negative intention group, 61.5% higher than positive intention male respondents, 19 to 21 years of age (0.108),
- (3) PMIL42 - male respondents, 16 to 18 years of age exhibit approximately 0.169 points higher propensity to apply than the negative intention group, 69.1% higher than positive intention male respondents, 19 to 21 years of age (0.052),

**Table 19. PPMM Estimates for Males, 16 to 18 Years of Age**

Number of Observations	6,487
Log Likelihood	-2592.73
Chi-Square	761.53

<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.1716
SEMA1721	-0.0004	-3.252	0.001	27.1261
DFJOB	-0.0281	-2.136	0.033	0.2166
DMILCOL	0.0640	2.402	0.016	0.0220
PMIL41	0.2797	15.998	0.000	0.0962
PMIL42	0.1685	11.094	0.000	0.1991
PMIL43	0.0622	4.293	0.000	0.3550
D16I	0.0894	4.584	0.000	0.3191
D17I	0.0462	2.373	0.018	0.3259
D1618NG	-0.0058	-0.270	0.787	0.2562
D1618G	0.0524	1.118	0.264	0.0130
DSOUTH	0.0153	1.030	0.303	0.3569
DNORTH	0.0226	1.438	0.150	0.2629
DWEST	0.0315	1.784	0.074	0.1821
DBLACK	0.0221	1.472	0.141	0.1131
DOTHER	0.0082	0.445	0.656	0.0908
DQ709P	-0.0550	-4.933	0.000	0.4098
Q700	0.0561	2.963	0.003	3.3502
Q700x2	-0.0059	-2.261	0.024	13.0139
Q713F	0.0524	3.933	0.000	13.1420
Q713Fx2	-0.0020	-3.964	0.000	180.9600
DRECR2	0.0671	5.713	0.000	0.2899
DASVAB	0.0543	2.610	0.009	0.0698
DAPPMET	0.1007	2.557	0.010	0.0127
DAPPMET	0.1257	1.810	0.070	0.0041
DAPPHS	0.0944	5.422	0.000	0.0901
CONSTANT	-0.8475	-8.789	0.000	1.0000

- (4) PMIL43 - male respondents, 16 to 18 years of age exhibit approximately 0.062 points higher propensity to apply than the negative intention group, 61.3% higher than PMIL43 male respondents, 19 to 21 years of age (0.024),
- (5) DRECR2 - male respondents, 16 to 18 years of age exhibit approximately 0.067 points higher propensity to apply than the non DRECR2 group, 54.7% higher than male respondents, 19 to 21 years of age (0.030),

**Table 20. PPMM Estimates for Males, 19 to 21 Years of Age**

Number of Observations	3,676
Log Likelihood	-748.09
Chi-Square	144.07

<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.0586
SEMA1721	-0.0000	-0.293	0.770	29.1736
DFJOB	-0.0184	-1.760	0.078	0.2576
DMILCOL	0.0490	1.179	0.238	0.0025
PMIL41	0.1076	6.716	0.000	0.0277
PMIL42	0.0521	4.334	0.000	0.1100
PMIL43	0.0241	2.507	0.012	0.3568
D190	0.0328	2.753	0.006	0.3722
D200	0.0115	0.876	0.381	0.2632
D1921I	0.0080	0.523	0.601	0.1400
DSOUTH	0.0320	2.500	0.012	0.3615
DNORTHC	0.0132	0.930	0.352	0.2568
DWEST	0.0377	2.520	0.012	0.1803
DBLACK	0.0026	0.208	0.835	0.0939
DOTHER	-0.0036	-0.248	0.804	0.0928
DQ709P	-0.0046	-0.486	0.627	0.2666
Q700	0.0103	0.642	0.521	3.5887
Q700x2	-0.0006	-0.288	0.773	14.4555
Q713F	0.0134	1.348	0.178	12.3478
Q713Fx2	-0.0006	-1.489	0.137	160.2637
DRECR2	0.0304	2.908	0.004	0.3464
DASVAB	0.0513	3.400	0.001	0.0686
DAPPMEP	0.0682	3.995	0.000	0.0342
DAPPMET	0.0597	1.806	0.071	0.0091
DAPPHS	0.0493	3.888	0.000	0.1180
CONSTANT	-0.3385	-4.885	0.000	1.0000

- (6) DASVAB - male respondents, 16 to 18 years of age exhibit approximately 0.054 points higher propensity to apply than the non DASVAB group, 5.5% lower than male respondents, 19 to 21 years of age (0.051),
- (7) DAPPMEP - male respondents, 16 to 18 years of age exhibit approximately 0.101 points higher propensity to apply than the non DAPPMEP group, 32.3% higher than male respondents, 19 to 21 years of age (0.068), and

- (8) DAPPHS - male respondents, 16 to 18 years of age exhibit approximately 0.094 points higher propensity to apply than the non DAPPHS group, 47.8% higher than male respondents, 19 to 21 years of age (0.049).

Tables 21 and 22 present the results of the equations for females, 16 to 18 years of age and 19 to 21 years of age, respectively. Results in the female equation are hampered by the small number of women in the survey samples and the resulting few numbers of females from the sample

**Table 21. PPM Estimates for Females, 16 to 18 Years of Age**

Number of Observations		3,328		
Log Likelihood		-441.41		
Chi-Square		128.27		
<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
APPLY				0.0351
SEMA1721	-0.0001	-0.767	0.443	26.9709
DFJOB	-0.0130	-1.282	0.200	0.1778
DMILCOL	0.0438	2.730	0.006	0.0084
PMIL41	0.0733	6.029	0.000	0.0273
PMIL42	0.0538	5.612	0.000	0.0721
PMIL43	0.0109	1.222	0.222	0.2376
D16I	0.0031	0.225	0.822	0.2900
D17I	0.0015	0.113	0.910	0.3349
D1618G	0.0002	0.013	0.990	0.3056
DSOUTH	-0.0035	-0.351	0.726	0.3599
DNORTHC	0.0037	0.358	0.720	0.2638
DWEST	-0.0061	-0.500	0.617	0.1830
DBLACK	0.0020	0.210	0.834	0.1223
DOTHER	0.0071	0.604	0.546	0.0932
DQ709P	-0.0183	-2.147	0.032	0.3165
Q700	0.0165	1.224	0.221	2.9399
Q700x2	-0.0018	-0.925	0.355	10.2130
Q713F	0.0026	0.317	0.751	13.0526
Q713Fx2	-0.0001	-0.411	0.681	178.5163
DRECR2	0.0119	1.372	0.170	0.1753
DASVAB	0.0353	3.502	0.000	0.0902
DAPPMEP	0.0104	0.288	0.773	0.0039
DAPPMET	0.0911	2.728	0.006	0.0018
DAPPHS	0.0249	2.035	0.042	0.0503
CONSTANT	-0.1733	-2.969	0.003	1.0000



who do eventually apply for military service. This is especially true for the older female group. Several variables had to be dropped from the estimation of the female equations due in large part to their small sample size. From the 16 to 18 year-old female equations, D1618G was dropped. The variables DMILCOL and DAPPMEP were dropped from the estimation of the 19 to 21 year-old equation.

**Table 22. PPMM Estimates for Females, 19 to 21 Years of Age**

		Number of Observations	2,110		
		Log Likelihood	-139.20		
		Chi-Square	98.79		
<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>	
APPLY				0.0178	
SEMA1721	-0.0001	-1.110	0.267	24.8960	
DFJOB	0.0064	0.846	0.398	0.2133	
PMIL41	0.0258	1.266	0.206	0.0069	
PMIL42	0.0114	0.853	0.394	0.0431	
PMIL43	0.0282	4.090	0.000	0.1948	
D190	0.0387	2.550	0.011	0.3741	
D200	0.0452	2.992	0.003	0.2646	
D1921I	0.0206	1.106	0.269	0.1194	
DSOUTH	0.0264	1.652	0.098	0.3765	
DNORTH	0.0262	1.596	0.111	0.2547	
DWEST	0.0385	2.315	0.021	0.1803	
DBLACK	0.0067	0.792	0.428	0.1070	
DOTHER	-0.0212	-1.276	0.202	0.1061	
DQ709P	0.0176	2.462	0.014	0.1975	
Q700	0.0103	0.754	0.451	3.1185	
Q700x2	-0.0006	-0.309	0.757	11.1999	
Q713F	-0.0030	-0.452	0.652	12.2646	
Q713Fx2	0.0001	0.230	0.818	158.6025	
DRECR2	0.0084	1.040	0.299	0.2083	
DASVAB	0.0080	0.602	0.547	0.0670	
DAPPMEP	0.0730	5.741	0.000	0.0102	
DAPPHS	0.0199	2.096	0.036	0.0731	
CONSTANT	-0.1537	-2.991	0.003	1.0000	

Coefficients vary widely between the two female age groups. The application rate is approximately 2 times higher for younger females (16 to 18) than it is for older females (19 to 21). In the younger female equation, PMIL41 and PMIL42 are significant at the 99% level of

confidence and the size of the coefficients follows the same descending pattern observed in the male equations. In the older female equation, only PMIL43 is significant at the 99% level of confidence. For younger females, DMILCOL appears to be important in determining their enlistment behavior. Variables for grades and father's education are statistically insignificant in both female equations, with the exception of DA in the younger female equation. Binary variables for contacting a recruiter and/or taking the ASVAB do not appear to be as important for females as was seen in the male equations. Not all recruiter and ASVAB variables are significant in the two equations, and the size of the coefficients of those significant at the 99% level of confidence is much smaller than those observed in the male equations.

### **Predicting Application Rates Using PPMM**

The male and female equations were then used to predict application (or enlistment) behavior both in and out-of-sample. These results are shown in Tables 23, 24, 25 and 26. The application rates presented in these tables represent the percent of respondents to a YATS survey who will apply for military service within 36 months of having been interviewed for the YATS. Both the actual application rates of YATS respondents and the predicted application rates display little annual movement. In-sample predictions were made for the years 1984 through 1987, and out-of-sample predictions were made for the years 1988 and 1989. The prediction out-of-sample for 1989 exhibited large errors, especially in the younger (16 to 18) male and female equations. This is largely a result of respondents to the 1989 YATS having had less time (only 24 months) in which to apply for military service when compared with earlier years within the sample. Thus, the actual application rate for 1989 represents application over only 24 months, while the PPMM prediction is an application rate over 36 months. This is especially true when considering 16 to 18 year-olds. Some respondents in the 16 to 18 year-old group must wait one year after the YATS before they are even eligible to apply for military service. Predictions for older females were once again hindered by the small numbers of older females surveyed in earlier years, and thus also the fewer number of them which eventually apply for military service.

**Table 23. Predicted Application Rate for Males, 16 to 18 Years of Age**

Year	Actual Apply	Predicted Apply	Difference	Percent Difference
1984	0.1846	0.1620	-0.0226	-12.24
1985	0.1654	0.1704	0.0050	3.02
1986	0.1725	0.1799	0.0074	4.29
1987	0.1607	0.1774	0.0167	10.39
1988	0.1596	0.1722	0.0126	7.89
1989 <sup>a</sup>	0.1390	0.1794	0.0404	29.06
Mean <sup>b</sup>	0.1686	0.1724	0.0038	2.25
Mean	0.1636	0.1736	0.0100	6.11

<sup>a</sup> 1989 YATS has been matched against fewer months of the MEPS files than the other years, thus the lower actual application rate

<sup>b</sup> Excludes 1989 in the calculation of the mean

**Table 24. Predicted Application Rate for Males, 19 to 21 Years of Age**

Year	Actual Apply	Predicted Apply	Difference	Percent Difference
1984	0.0562	0.0614	0.0052	9.25
1985	0.0612	0.0555	-0.0057	-9.31
1986	0.0630	0.0569	-0.0061	-9.68
1987	0.0540	0.0602	0.0062	11.48
1988	0.0634	0.0615	-0.0019	-3.00
1989 <sup>a</sup>	0.0659	0.0598	-0.0061	-9.26
Mean <sup>b</sup>	0.0596	0.0591	-0.0005	-0.84
Mean	0.0606	0.0592	-0.0014	-2.31

<sup>a</sup> 1989 YATS has been matched against fewer months of the MEPS files than the other years, thus the lower actual application rate

<sup>b</sup> Excludes 1989 in the calculation of the mean

**Table 25. Predicted Application Rate for Females, 16 to 18 Years of Age**

Year	Actual Apply	Predicted Apply	Difference	Percent Difference
1984	0.0306	0.0354	0.0048	15.69
1985	0.0405	0.0330	-0.0075	-18.52
1986	0.0365	0.0351	-0.0014	-3.84
1987	0.0334	0.0370	0.0036	10.78
1988	0.0447	0.0362	-0.0085	-19.02
1989 <sup>a</sup>	0.0296	0.0355	0.0059	19.93
Mean <sup>b</sup>	0.0371	0.0353	-0.0018	-4.85
Mean	0.0359	0.0354	0.0005	1.39

<sup>a</sup> 1989 YATS has been matched against fewer months of the MEPS files than the other years, thus the lower actual application rate

<sup>b</sup> Excludes 1989 in the calculation of the mean

**Table 26. Predicted Application Rate for Females, 19 to 21 Years of Age**

Year	Actual Apply	Predicted Apply	Difference	Percent Difference
1984	0.0287	0.0216	-0.0071	-24.74
1985	0.0105	0.0150	0.0045	42.86
1986	0.0173	0.0169	-0.0004	-2.31
1987	0.0102	0.0162	0.0060	58.82
1988	0.0180	0.0191	0.0011	6.11
1989 <sup>a</sup>	0.0131	0.0196	0.0065	49.62
Mean <sup>b</sup>	0.0169	0.0178	0.0009	5.33
Mean	0.0163	0.0181	0.0018	11.04

<sup>a</sup> 1989 YATS has been matched against fewer months of the MEPS files than the other years, thus the lower actual application rate

<sup>b</sup> Excludes 1989 in the calculation of the mean

## VI. MULTIPLE DECISION MODELS

Conceptually, the decision to join the military can be considered an occupational decision, i.e., choosing between the military versus numerous other non-military job opportunities. However, some choices for the military are more like a paid preschooling decision, e.g., a means of obtaining partial funding for college (Army College Fund). The choice which is most unlike going to work or joining the military is the decision to attend school. One way of modeling this trichotomy is through the use of a multinomial logit estimator which allows for multiple zero/one type decisions where more than two alternatives exist. Multinomial logit will provide three sets of coefficients for each possible two-way decision, i.e., military versus school, military versus work, and school versus work.

For the multinomial logit estimation, the dependent variable,  $y$ , will have three possible outcomes, 1 for choosing military, 2 for choosing work, and 3 for choosing school. The values of  $y$  are assumed to be "unordered" in the multinomial logit estimation. That is, the value assigned to the dependent variable are arbitrary in the sense that  $1 < 2 < 3$  does not imply that outcome 1 (military) is less than outcome 2 (work) is less than outcome 3 (school). The explanatory variables of the estimation are defined by  $X$ . In the multinomial logit estimation, a set of coefficients  $B^{(1)}$ ,  $B^{(2)}$ , and  $B^{(3)}$  are estimated for each possible outcome of the dependent variable in the form:

$$P(y=1) = \frac{e^{XB^{(1)}}}{e^{XB^{(1)}} + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (1)$$

$$P(y=2) = \frac{e^{XB^{(2)}}}{e^{XB^{(1)}} + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (2)$$

$$P(y=3) = \frac{e^{XB^{(3)}}}{e^{XB^{(1)}} + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (3)$$

The model as specified in Equations 1, 2, and 3, however, is unidentified in the sense that there is more than one solution to the coefficients that leads to the same probabilities for  $y = 1$ ,  $y = 2$ , and  $y = 3$ . To identify the model, one of the set of coefficients is arbitrarily set to zero. For example, if  $B^{(1)}$  is set to zero, the remaining coefficients,  $B^{(2)}$  and  $B^{(3)}$ , would measure the change relative to the  $y = 1$  group (military). Setting  $B^{(1)} = 0$ , the equations for each outcome can then be expressed as:

$$P(y=1) = \frac{1}{1 + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (4)$$

$$P(y=2) = \frac{e^{XB^{(2)}}}{1 + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (5)$$

$$P(y=3) = \frac{e^{XB^{(3)}}}{1 + e^{XB^{(2)}} + e^{XB^{(3)}}} \quad (6)$$

The coefficients of the multinomial logit are not strictly interpretable as the quantitative measure of a change in  $y$  for a given change in an explanatory variable. To simplify the quantitative interpretation of the multinomial coefficients, the relative probability of  $y = 2$  to the base category ( $y = 1$ ) can be determined by the ratio:

$$\frac{P(y=2)}{P(y=1)} = e^{XB^{(2)}} \quad (7)$$

This ratio can be interpreted as the relative risk of  $y = 2$  to the base category. Assuming  $X$  and  $B^{(2)}$  are vectors equal to  $x_1, x_2, \dots, x_k$  and  $B_1^{(2)}, B_2^{(2)}, \dots, B_k^{(2)}$ , respectively, the ratio of the relative risk for a one unit change in  $x_i$  is then equal to:

$$e^{B_i^{(2)}} = \frac{e^{B_1^{(2)}x_1 + \dots + B_i^{(2)}(x_i+1) + \dots + B_k^{(2)}x_k}}{e^{B_1^{(2)}x_1 + \dots + B_i^{(2)}x_i + \dots + B_k^{(2)}x_k}} \quad (8)$$

Thus, the exponential value of a coefficient is the relative risk ratio (RRR) for a one unit change in the corresponding variable  $x_i$ , understanding that the risk is being measured as the risk of the category relative to the base category ( $y = 1$ ).

For example, if school ( $y = 2$ ) were to be defined as the base category, and the RRR for the binary variable for DBLACK were 1.25 for military ( $y = 1$ ), this would imply that the probability of choosing the military over school would increase by 25% if the respondent is black (DBLACK=1), regardless of the values of the other explanatory variables. Given the same base category school ( $y = 3$ ), an RRR for the variable D1618G of 0.17 implies that the probability of choosing military over school decreases by 83% if the respondent has graduated from high school (D1618G=1). For ease of interpretation, RRR values will be presented in the multinomial logit estimation results presented in this section.

### Multinomial Logit Applied to Work, School, and Military

*Question 517* was used to construct the dependent variable for the multinomial logit estimation. *Q517* prompts the respondent for what he plans to be doing one year from the survey date, or after graduation from high school. The respondent must choose from a set of options. The dependent variable has three possible outcomes with the results providing information concerning each of the two-way choices, e.g., military or school, but not work. Table 27 shows how the dependent variable for the estimation was constructed from *Q517*. Since what the respondent actually did can not be determined with the exception of those who opted to not join the military, responses to *Q517* are being used as an indication of the actual selections of the respondents.

Table 28 shows that the greatest proportion of the respondents select going to school over going to work or joining the military. The percentages fluctuate by age. When these proportions are viewed by age group (Table 29), respondents ages 16 through 19 tend to select going to school over going to work or joining the military. Older respondents, those age 20 and

**Table 27. Dependent Variable - Multinomial Logit**

Value	Description
1 - Military	<i>Q517</i> is equal to "5" (serving in the military)
2 - Work	<i>Q517</i> is equal to "3" (working full-time) or "6" (being a full-time homemaker) or if <i>Q517</i> is equal to "4" (working part-time) and <i>Q514</i> is equal to "3" or "4" (probably not or definitely not going to college)
3 - School	<i>Q517</i> is equal to "1" (going to school full-time) or "2" (going to school part-time) or if <i>Q517</i> is equal to "4" (working part-time) and <i>Q514</i> is equal to "1" or "2" (definitely or probably going to college)

21, tend to select going to work over going to school or joining the military. In fact, the expressed desire to enter the military significantly declined in the age groups 19 to 21. This suggests that if the commitment has not been made prior to the age of 19, the likelihood of engendering a positive propensity rapidly declines. This suggests that the 16 to 18 year-old group is possibly the best group at which recruiting efforts should be directed since they reflect the highest propensities for military service.

**Table 28. Military v. Work v. School**

Year	Percent Military	Percent Work	Percent School
1984	4.79	37.95	57.26
1985	5.15	34.67	60.18
1986	5.60	33.96	60.44
1987	5.96	31.72	62.32
1988	5.80	29.64	64.56
1989	6.18	28.99	64.83

Three equations will be estimated using multinomial logit. An equation for young males (16 to 18 years old), older males (19 to 21 years old), and young females (16 to 18 years old). Due to the small number of older females (19 to 21 years old) surveyed, an equation for this group could not be estimated using multinomial logit. Respondents with a negative SIGNDT

are excluded from the estimation sample for all groups. These respondents are excluded from the estimation because, having already made a formal commitment to the military (hence, decided military), they could potentially bias the estimation results.

**Table 29. Military v. Work v. School by Age**

Age	Percent Military	Percent Work	Percent School
16	10.04	17.22	72.74
17	8.22	19.78	72.00
18	5.10	26.57	68.33
19	2.53	36.64	60.83
20	2.15	58.06	39.79
21	1.51	67.49	31.00

### **Multinomial Logit Results**

The explanatory variables used in the estimation equations are all binary variables with the exception of population (SEMA1721). Responses to Q700 (grades) and Q713F (father's education) have been used to construct binary variables for the estimation equations. Table 30 shows the new binary variables created for grades and father's education. A binary variable for taking or planning to take Calculus was also created. A detailed description of how these variables were constructed is provided in Appendix C. All other variables in the equation are constructed in the same manner as those used in the PPMM estimation in Section V. Tables 31 and 32 present two sets of relative risk ratios (RRRs) for the young male (16 to 18 year-old) equations for each of the unique two-way choices: Table 31 presents military/school and work/school choices (school is base category), and Table 32 presents military/work and school/work choices (work is base category).

For the younger males, the military versus school and work versus school decisions seem to be quite similar (Table 31). For younger males, respondents with higher grades (DA, DAB, DB, and DBC) show lower propensities to choose the military or work over going to school.

**Table 30. Binary Variables for Multinomial Logit Estimation**

Variable Name	Description
DQ708P	Taken or plan to take calculus
DA	Received mostly A's in high school
DAB	Received mostly A's and B's in high school
DB	Received mostly B's in high school
DBC	Received mostly B's and C's in high school
DC	Received mostly C's in high school
DDF	Received mostly D's or F's in high school
DFGR12	Father did not complete high school
DFGR3	Father has a high school diploma
DFGR4	Father has some college
DFGR5	Father has a college degree or greater

Respondents whose fathers received less education than respondents whose fathers held at least a college degree, showed higher propensities to choose the military or going to work over going to school. The lower the level of the father's education, the greater the increase in the propensity to choose the military or work over school. Respondents still in high school showed lower propensities to choose work over going to school. Black respondents showed higher propensities to choose military over school (compared to Caucasian, non-Hispanic respondents), but lower propensities to choose work over school. Having contacted a recruiter and taken the ASVAB at a MEPS, METS, or in high school, or only contacting a recruiter, increased the propensity to choose military over going to school. Respondents who had taken the ASVAB in high school also showed lower propensity for choosing work over school.

For younger males comparing the military with going to work (Table 32), having contacted a recruiter or taken the ASVAB at a MEPS or in high school increased their propensity for choosing the military over work. Respondents with average grades (DBC) showed high propensity to choose the military over work. Blacks and other minorities showed higher propensities than Caucasians to select the military over going to work. Having contacted

**Table 31. Young Males - Military/School - Work/School**

Number of Observations

6,382

<u>Name</u>	<u>Military/School</u>		<u>Work/School</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	1.0006	0.648	0.9986	-1.988
D16I	1.4206	2.103	0.6346	-3.739
D17I	0.9308	-0.430	0.6964	-3.047
D1618G	0.1643	-7.947	0.6620	-3.385
D1618NG	0.5226	-1.333	1.0694	0.235
DSOUTH	1.1246	0.794	0.9804	-0.209
DNORTHC	1.4828	2.584	0.9056	-0.989
DWEST	1.2048	1.047	1.0100	0.086
DBLACK	1.8338	4.760	0.6492	-3.772
DOTHER	1.2674	1.414	0.7540	-2.239
DQ708P	0.6880	-2.795	0.5268	-6.080
DQ709P	0.9247	-0.693	0.6292	-5.469
DA	0.1878	-4.871	0.1723	-8.062
DAB	0.3026	-5.243	0.2339	-9.137
DB	0.3507	-4.729	0.2582	-8.773
DBC	0.6956	-1.786	0.4516	-5.518
DC	0.7680	-1.218	0.6806	-2.547
DFGR12	4.0971	8.499	4.4579	13.224
DFGR3	2.6441	6.512	2.8890	10.876
DFGR4	2.0541	3.859	1.6166	3.654
DRECR2	2.7409	9.348	0.9120	-1.164
DASVAB	0.8084	-0.813	0.8729	-1.031
DAPPMEP	10.8035	7.823	1.0864	0.265
DAPPMET	4.8621	2.300	1.4002	0.677
DAPPHS	2.6912	5.820	0.7200	-2.545

a recruiter or taken the ASVAB at a MEPS, METS, or in high school greatly increased the propensity to choose joining the military over going to work.

For older males, the military versus school and work versus school decisions (Table 33) showed more differences than were seen in the younger male equation. For older males, respondents with higher grades (DA, DAB, and DB) showed lower propensities to choose work over going to school. Grades do not appear to be a factor in the military versus school decision for older males. As with the younger males, respondents whose father received less education than respondents whose father held at least a college degree, showed higher propensities to choose the military or going to work. The lower the level of the father's education, the greater

**Table 32. Young Males - Military/Work - School/Work**

Number of Observations

6,382

<u>Name</u>	<u>Military/Work</u>		<u>School/Work</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	1.0021	1.962	1.0015	1.988
D16I	2.2388	4.561	1.5759	3.739
D17I	1.3364	1.654	1.4359	3.047
D1618G	0.2482	-5.942	1.5105	3.385
D1618NG	0.4887	-1.404	0.9351	-0.235
DSOUTH	1.1471	0.856	1.0200	0.209
DNORTH	1.6373	2.974	1.1042	0.989
DWEST	1.1928	0.910	0.9901	-0.086
DBLACK	2.8249	6.947	1.5405	3.772
DOTHER	1.6809	2.769	1.3262	2.239
DQ708P	1.3059	1.691	1.8981	6.080
DQ709P	1.4696	3.021	1.5893	5.469
DA	1.0900	0.229	5.8027	8.062
DAB	1.2939	1.076	4.2757	9.137
DB	1.3586	1.329	3.8736	8.773
DBC	1.5402	2.095	2.2142	5.518
DC	1.1285	0.553	1.4693	2.547
DFGR12	0.9191	-0.456	0.2243	-13.224
DFGR3	0.9152	-0.523	0.3461	-10.876
DFGR4	1.2707	1.112	0.6186	-3.654
DRECR2	3.0053	9.229	1.0965	1.164
DASVAB	0.9261	-0.279	1.1456	1.031
DAPPMEP	9.9445	6.648	0.9205	-0.265
DAPPMET	3.4726	1.761	0.7142	-0.677
DAPPHS	3.7377	6.847	1.3888	2.545

the increase in the propensity to choose the military or work over school. Nineteen and twenty year-old respondents that had graduated from high school showed lower propensities than twenty-one year-old high school graduates to choose work over going to school. Black respondents showed higher propensities to choose military over school (compared to Caucasian, non-Hispanic respondents), but lower propensities to choose work over school. Having contacted a recruiter or taken the ASVAB at a MEPS greatly increased the propensity to choose military over school.

For older males, comparing the military with going to work (Table 34), having contacted a recruiter or taken the ASVAB at a MEPS increased propensity for choosing the military

**Table 33. Older Males - Military/School - Work/School**

Number of Observations

3,623

<u>Name</u>	<u>Military/School</u>		<u>Work/School</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	0.9999	-0.040	0.9986	-1.976
D190	0.8407	-0.376	0.2422	-13.736
D200	0.7117	-0.644	0.5257	-5.907
D1921I	4.9157	3.403	0.7416	-2.114
DSOUTH	1.9519	1.693	0.7582	-2.582
DNORTHC	1.0524	0.113	0.7203	-2.910
DWEST	0.6891	-0.656	0.5176	-5.169
DBLACK	2.7735	3.246	0.7156	-2.503
DOTHER	0.9099	-0.189	0.6937	-2.715
DQ708P	1.0984	0.227	0.6169	-3.999
DQ709P	0.6018	-1.519	0.6000	-5.442
DA	0.8149	-0.176	0.3478	-4.017
DAB	2.5346	1.284	0.5541	-2.851
DB	1.0618	0.080	0.5384	-3.120
DBC	1.9288	0.964	0.7647	-1.389
DC	1.4110	0.488	0.7390	-1.510
DFGR12	3.8847	2.433	3.4105	9.794
DFGR3	4.0225	2.696	2.6112	8.929
DFGR4	1.9053	0.986	1.3106	1.933
DRECR2	3.6981	3.755	0.8536	-1.824
DASVAB	0.4009	-0.761	1.0003	0.002
DAPPMET	9.5321	4.514	1.0221	0.102
DAPPMET	4.9817	1.487	1.1610	0.360
DAPPHS	1.3217	0.479	0.9834	-0.136

over work. Respondents with better than average grades (DAB) showed higher propensity to choose the military over work. Blacks and other minorities showed higher propensities than Caucasians to select the military over going to work. Respondents who were 19 years old and had graduated from high school showed higher propensity to choose the military over going to work when compared with 21 year-old high school graduates. Also, older males who had not completed high school showed higher propensity to choose the military over going to work compared to 21 year-old high school graduates.

Younger females choosing between the military and school and work and school (Table 35) showed results similar to those of the young males. Father's education and grades were statistically significant in the decision to choose work over school. Respondents with higher

**Table 34. Older Males - Military/Work - School/Work**

Number of Observations

3,623

<u>Name</u>	<u>Military/Work</u>		<u>School/Work</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	1.0013	0.487	1.0014	1.976
D190	3.4703	2.713	4.1281	13.736
D200	1.3539	0.578	1.9022	5.907
D1921I	6.6287	4.118	1.3485	2.114
DSOUTH	2.5745	2.414	1.3190	2.582
DNORTHC	1.4611	0.846	1.3884	2.910
DWEST	1.3314	0.506	1.9321	5.169
DBLACK	3.8760	4.367	1.3975	2.503
DOTHER	1.3117	0.545	1.4416	2.715
DQ708P	1.7807	1.391	1.6211	3.999
DQ709P	1.0030	0.009	1.6666	5.442
DA	2.3431	0.735	2.8754	4.017
DAB	4.5739	2.130	1.8046	2.851
DB	1.9722	0.924	1.8573	3.120
DBC	2.5222	1.382	1.3077	1.389
DC	1.9094	0.934	1.3532	1.510
DFGR12	1.1390	0.233	0.2932	-9.794
DFGR3	1.5405	0.833	0.3830	-8.929
DFGR4	1.4537	0.569	0.7630	-1.933
DRECR2	4.3325	4.234	1.1716	1.824
DASVAB	0.4008	-0.762	0.9997	-0.002
DAPPMET	9.3257	4.585	0.9783	-0.102
DAPPMET	4.2909	1.383	0.8613	-0.360
DAPPHS	1.3440	0.510	1.0169	0.136

grades and fathers with less education exhibited lower propensities to choose going to work over going to school. Blacks showed higher propensity than Caucasians to choose the military over school. Sixteen and seventeen year-old respondents still in high school showed lower propensities to choose work over school than eighteen year-olds still in high school. As with the males, female respondents who had contacted a recruiter or taken the ASVAB at a MEPS or in high school displayed higher propensity to choose the military over going to school. For females, comparing the military with going to work (Table 36), having contacted a recruiter or taken the ASVAB at a MEPS or in high school increased propensity for choosing the military over work. Black female respondents showed higher propensities than Caucasians to select the military over going to work.

**Table 35. Young Females - Military/School - Work/School**

Number of Observations 3,299

<u>Name</u>	<u>Military/School</u>		<u>Work/School</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	0.9912	-1.856	0.9971	-2.656
D16I	1.9497	1.373	0.4641	-4.011
D16I	0.9187	-0.173	0.4874	-3.860
D1618G	0.2731	-2.296	0.7180	-1.783
D1618NG	1.6983	0.674	0.8032	-0.548
DSOUTH	1.2881	0.734	1.0523	0.378
DNORTHC	1.0390	0.100	0.9411	-0.427
DWEST	0.9483	-0.116	1.2937	1.584
DBLACK	3.5820	4.898	0.7458	-1.930
DOTHER	1.0982	0.194	0.7968	-1.327
DQ708P	0.6383	-1.470	0.5194	-4.125
DQ709P	1.4784	1.508	0.8428	-1.362
DA	0.2629	-1.469	0.2117	-4.456
DAB	0.6073	-0.710	0.3515	-3.538
DB	0.8145	-0.291	0.3617	-3.417
DBC	0.9365	-0.094	0.8328	-0.630
DC	1.8587	0.872	1.1614	0.488
DFGR12	3.1402	2.984	4.0886	8.690
DFGR3	1.8698	1.716	2.7433	7.057
DFGR4	0.8770	-0.249	1.5706	2.405
DRECR2	5.0265	6.331	1.1195	0.906
DASVAB	0.2195	-1.331	0.8283	-1.110
DAPPMEP	12.1184	3.000	0.2354	-1.606
DAPPHS	6.6490	4.986	0.7867	-1.052

The multinomial logit estimation results tend to suggest that, in general, the decision to join the military is more similar to the decision to go to work than to attend school. Respondents' grades and their father's education were in many cases important factors in influencing the outcome of the military, work, school decision. This was especially true for younger respondents. The estimations also tended to suggest that blacks were more likely than non-Hispanic Caucasians to choose to join the military over either going to work or attending school.

**Table 36. Young Females - Military/Work - School/Work**

Number of Observations 3,299

<u>Name</u>	<u>Military/Work</u>		<u>School/Work</u>	
	<u>RRR</u>	<u>t-statistic</u>	<u>RRR</u>	<u>t-statistic</u>
SEMA1721	0.9941	-1.232	1.0029	2.656
D16I	4.2009	2.886	2.1546	4.011
D17I	1.8850	1.270	2.0518	3.860
D1618G	0.3803	-1.685	1.3928	1.783
D1618NG	2.1144	0.916	1.2450	0.548
DSOUTH	1.2241	0.564	0.9503	-0.378
DNORTHC	1.1040	0.249	1.0626	0.427
DWEST	0.7330	-0.657	0.7730	-1.584
DBLACK	4.8028	5.553	1.3408	1.930
DOTHER	1.3782	0.646	1.2550	1.327
DQ708P	1.2289	0.623	1.9252	4.125
DQ709P	1.7541	2.045	1.1865	1.362
DA	1.2420	0.231	4.7238	4.456
DAB	1.7278	0.759	2.8449	3.538
DB	2.2520	1.123	2.7648	3.417
DBC	1.1246	0.165	1.2008	0.630
DC	1.6004	0.648	0.8610	-0.488
DFGR12	0.7680	-0.652	0.2446	-8.690
DFGR3	0.6816	-0.997	0.3645	-7.057
DFGR4	0.5584	-1.058	0.6367	-2.405
DRECR2	4.4898	5.582	0.8932	-0.906
DASVAB	0.2650	-1.158	1.2073	1.110
DAPPMEP	51.4839	3.426	4.2484	1.606
DAPPHS	8.4517	5.125	1.2711	1.052

## **VII. MODEL METHODOLOGY AND PROJECTION**

In this section, a methodology is described for extrapolating the predictions from a logit model (such as those presented in Tables 19, 20, 21, and 22) to the population of potential applicants. This technique uses aggregate rate projections from the logit model combined with a transitional regression equation to project overall application rates for males and females in the 17 to 21 year age range. The methodology requires the completion of four processing steps: (1) select the respondents to be included in the projection, (2) estimate logit application equation for projection, (3) apply the logit application equation from the YATS respondents, (4) compute application rates by FY from the equation results, and (5) adjust the FY annual rates using a transitional equation. Each of these steps will be described below, and the results presented for projecting FY application rates from 1986 through 1991.

All results from the methodology are presented at the aggregate 17 to 21 (or 18 to 21) age level. Initial attempts to project rates by age cohort and to extend the model to higher age groups have been hampered by the small number of YATS respondents who apply during the relevant time frame. As can be seen in Table 37, most age specific male cohorts contain fewer than 100 YATS respondents who apply within 0 to 35 months of taking the YATS. These numbers are far smaller for females of similar age groups. For projection purposes, these small numbers cause further problems because of the temporal nature of the applicants' arrivals. Arrivals must be projected to at least the single FY level as opposed to the 3 year horizon used in the earlier equations. Attempts to extrapolate to national application rates from such small numbers of positive application outcomes pose both theoretical and empirical problems.

### **Respondent Selection**

Two factors are considered in including YATS respondents in the projections process — age and military status when the YATS was administered. The projection model projects application rates by FY for 17 through 21 year-olds. As described in the following section, this is the age at the time of application and corresponds to a decision an average of two years after the date of the survey. This restricts the age at the time of the survey to respondents between

**Table 37. YATS Respondents Applying Within Three Years of YATS**

Year	Age when YATS taken					Total
	16	17	18	19	20	
84	88	136	87	36	19	366
85	129	103	73	34	18	357
86	117	104	75	34	19	349
87	140	136	52	37	12	377
88	107	131	72	29	14	353
89	100	136	83	45	23	387
Total	681	746	442	215	105	2,189

15 and 19 inclusive. Those outside this age range at the time of the survey are excluded from all proceeding analysis. Obviously 15 year-olds are not available on the YATS and a correction will be made for this in the transitional equation.

The projection model is to be applied to a military eligible population, and those respondents who have already applied for military service must be excluded. This requires that the YATS file be merged against MEPS files from at least the three years prior to the YATS. Matching can be performed only for respondents from whom a Social Security number was obtained, and those without Social Security numbers are also dropped. Those YATS respondents who match a MEPS application or enlistment record are also dropped from consideration in all of the following steps: Finally, any respondent whose responses cannot be used to generate the variables used in the logit equation described below must also be excluded.

### **Projection Model Equation**

The second step toward building a projection model involves estimating a FY specific logit mode based on the results presented earlier. This model differs from the prior models primarily in the formulation of its independent variable. Rather than assign a positive outcome (1.0) to anyone who applies within 3 years of taking the YATS, the specific FY to be projected is used. Since the YATS requires several months of processing to be completed, it is the FY

which follows the year in which the YATS is administered which is most appropriate. For example, the 1989 YATS is completed toward the end of calendar year '89 and the start of FY 90; the model will project FY 91 applicants from this YATS. In this way, only respondents who apply in the complete FY following the end of a YATS survey are assigned a positive response; all others are assigned a zero. Several attempts were made to utilize models with a 3 year definition of applicants combined with separate models for distributing the arrival of applicants over FY's, quarters, or months. The primary obstacle in these models is shifting the age distribution of applicants, and these models consistently performed poorly at projecting FY application rates.

This definition of an applicant significantly reduces the number of positive dependent variables and requires some simplification of the models discussed earlier. In addition, merging all age groups and both genders into the same equation allowed the model to be extended to females and the higher age groups. Earlier attempts to project with separate gender and 19-21 age equations produced poor projections for these groups. The simplification of the earlier models primarily involved the combination of the indicator variables and dropping some variables. Table 38 defines those variables which have not appeared in a prior model. The gender/age indicators use the respondent's age at the time of the YATS in the construction of the variable; these ages would be 2 years larger on average at the time of application.

The projection model differs from the earlier models in one final and important area. It includes a proxy variable for the military's demand for applicants. The application rate demonstrates temporal patterns which cannot be completely explained by labor supply and attitude conditions and are hypothesized to result from demand conditions. Using the application rate as the dependent variable in current and prior models is an attempt to reduce the impact of demand factors and to model a "free-flow" variable which is purely a decision for an individual youth. However, even with applicants, recruiters have considerable power in determining whether the application process is completed. Furthermore, the position of the recruiter is clearly related to the current accession needs of his/her respective service. This influence is captured in reduced form in the current model by including a proxy for total military FY accession goals. To be applicable to individual decisions, these goals should be divided by the eligible population of applicants so that they indicate the probability of selection for a typical

**Table 38. New Variables in the Projection Model**

<b>Variable Name</b>	<b>Variable Description</b>	<b>YATS Questions</b>
<b>APP_FY1</b>	Binary: 1 = Applied in the FY following the FY in which the YATS was completed, 0 = Otherwise	N.A.
<b>DMAL16</b>	Binary: 1 = Respondent is male and was 16 when the YATS was taken, 0 = Otherwise	Q402 is equal to "1" and Q403 is equal to "16"
<b>DMAL17</b>	Binary: 1 = Respondent is male and was 17 when the YATS was taken, 0 = Otherwise	Q402 is equal to "1" and Q403 is equal to "17"
<b>DMAL18</b>	Binary: 1 = Respondent is male and was 18 when the YATS was taken, 0 = Otherwise	Q402 is equal to "1" and Q403 is equal to "18"
<b>DFEM16</b>	Binary: 1 = Respondent is female and was 16 when the YATS was taken, 0 = Otherwise	Q402 is equal to "2" and Q403 is equal to "16"
<b>DFEM17</b>	Binary: 1 = Respondent is female and was 17 when the YATS was taken, 0 = Otherwise	Q402 is equal to "2" and Q403 is equal to "17"
<b>DFEM18</b>	Binary: 1 = Respondent is female and was 18 when the YATS was taken, 0 = Otherwise	Q402 is equal to "2" and Q403 is equal to "18"
<b>DFEM19</b>	Binary: 1 = Respondent is female and was 19 when the YATS was taken, 0 = Otherwise	Q402 is equal to "2" and Q403 is equal to "19"
<b>DQ709</b>	Binary: 1 = has taken high school physics, 0 = Otherwise	Q709 is equal to "1"
<b>DRECR</b>	Binary: 1 = has talked to military recruiter, 0 = Otherwise	Q628 is equal to "1"
<b>DASVAB</b>	Binary: 1 = has taken the ASVAB but not talked to a recruiter, 0 = Otherwise	Q645 is equal to "1" and Q628 is equal to "2"
<b>CON_RAT1</b>	Total military contracts divided by the 17 to 21 year-old population. A proxy for FY accession goals relative to the eligible population size.	N.A.

individual. This variable is proxied by the actual number of military contracts in a fiscal year divided by the size of the 17 to 21 year-old population. It should be noted that the population is estimated for the end of a calendar year, accessions over the complete FY.

The results of estimating the projection model logit equation on the sample described in the prior section are presented in Table 39. This model was developed using the WINT weights discussed earlier and, like the prior models, was estimated over the 1984 through 1987 YATS.

**Table 39. Projection Model Estimate  
Males and Females, 18 to 21 During Application FY**

Variable Name	Coefficient	t-statistic	Prob >  t
DMAL16	1.4199	6.260	0.000
DMAL17	0.8077	3.473	0.000
DMAL18	0.3091	1.215	0.224
DFEM16	0.0244	0.083	0.934
DFEM17	-0.5966	-1.853	0.064
DFEM18	-0.5848	-1.706	0.088
DFEM19	-0.3145	-0.957	0.339
Q700	0.5954	2.816	0.005
Q700X2	-0.0657	-2.263	0.024
Q713F	0.2689	1.910	0.056
Q713FX2	-0.0107	-2.000	0.045
DQ709	-0.5095	-3.128	0.002
DFJOB	-0.4128	-2.677	0.007
SEMA1721	-0.0378	-1.141	0.254
DRECR	0.3180	2.675	0.007
DASVAB	0.5299	2.658	0.008
DMILCOL	0.1555	0.788	0.430
PMIL41	1.2872	6.138	0.000
PMIL42	1.0343	6.500	0.000
PMIL43	0.4265	2.833	0.005
CON_RAT1	126.6333	1.537	0.124
CONSTANT	-8.4140	-6.330	0.000

### **Application of the Projection Model Equation**

The third step involves applying the logit enlistment equation from Table 39 to all the sample of YATS respondents described in the respondent selection section. As noted, these respondents will have been between the ages of 16 and 19 inclusive when the interview was performed. The equation is applied by multiplying each coefficient by its associated YATS

response or derived variable. The resulting values are summed to produce a linear response  $L$ . This linear response is then evaluated using a logit transformation to produce the probability that the individual will apply during the first full FY following the completion of the YATS. For example, the calendar year 89 YATS will be used to project the FY 91 applicants. Thus, the 16 through 19 year-olds will have aged to 17 through 21 year-olds at their time of application. The logit transformation which produces a probability can be expressed in terms of the linear response as seen in Equation 9. At the completion of this step, each selected respondent has been assigned an individual probability of applying during the designated fiscal year.

$$probability = \frac{1}{1 + e^{-L}} \quad (9)$$

### Computing Fiscal Year Average Probabilities

The fourth step in the methodology requires evaluating the individual application probabilities to produce FY weighted averages of application probabilities by gender for youth age 18 to 21. For each year in which a YATS survey is available, the weighted average of the individual probabilities indicates the application rate for the cohort during the target FY. The individual probabilities are weighted using the same weight as was employed in the development of the application model (WINT). Equation 10, demonstrates the computation of these weighted averages for males from a single YATS survey. The weighted rate for females is computed in exactly the same manner. The result of this step is an expected application probability for any youth age 18 to 21 and of the appropriate gender. This expected probability may be interpreted as a projected application rate. When using a single YATS survey, the average application probability projection is for a single FY; when several YATS surveys are used, the collection of average probabilities forms a time-series.

$$Rate_{males} = \frac{\sum_{i \in sample} P_i W_i}{\sum_{i \in sample} W_i} \quad (10)$$

where:

- Rate<sub>males</sub>* is the weighted application probability over all of the males in the selected sample. (ie. The expected application probability.)
- P<sub>i</sub>* is the probability (from the logit evaluation) of an individual applying in the target FY.
- W<sub>i</sub>* is the WINT weight for an individual.
- sample* is the sample defined in the section on selecting a sample and is also restricted to males.

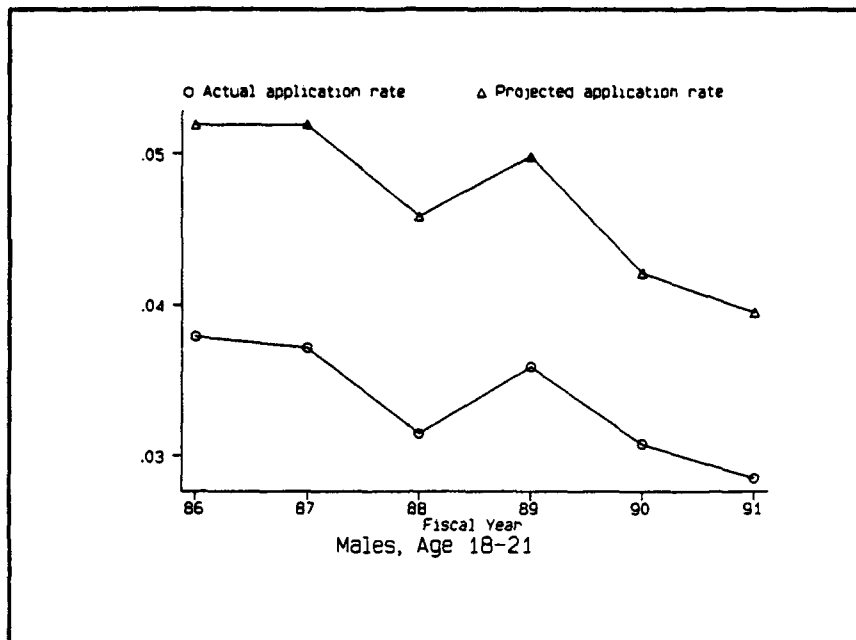
When this probability assignment and weighted summation process is performed for each of the YATS years 1984 through 1989, a series of projected average probabilities is obtained for males age 18 to 21 and for females age 18 to 21. These projections cover the fiscal years 1986 through 1991. Figure 1 presents the actual and projected application rates for males while Figure 2 presents the same information for females. It should be noted that the projections for 1990 and 1991 are truly out-of-sample projections as the equation from Table 39 was only estimated over the 1984 through 1987 YATS. The 1988 and 1989 YATS which form the basis of the 1990 and 1991 FY estimates were not used to develop the projection model. The actual rates are taken from MEPCOM data with the gender specific youth population between ages 18 and 21 serving as the denominator.

As can be seen in both cases, the actual rates and the weighted projections from the YATS track extremely closely, but the actual rates are substantially lower than the weighted projected rates. This observation is consistent with the concerns expressed earlier about the potential for those most interested in the military self selecting into the sample at a higher rate

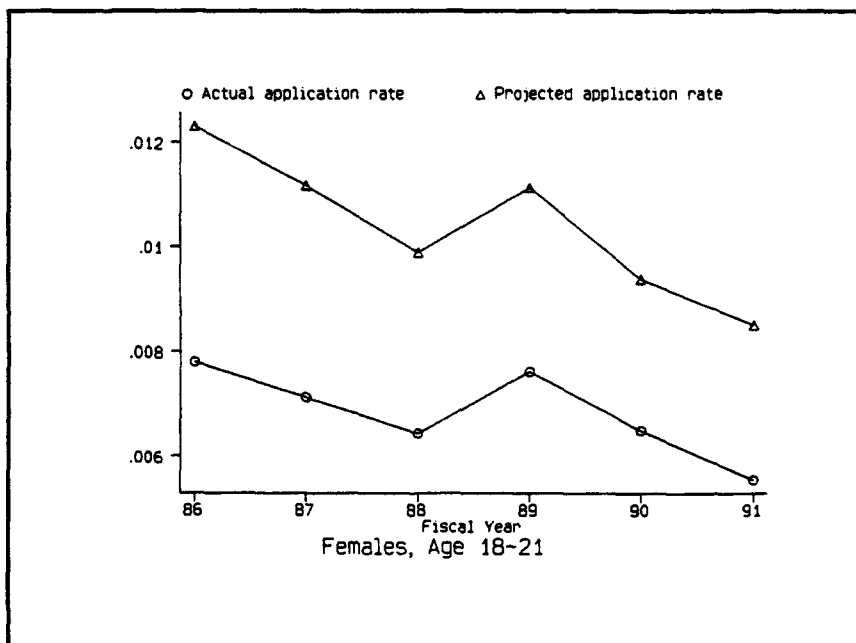
than those who are uninterested. Put another way, sampled youth who are uninterested in the military may be more likely to refuse to participate in the YATS.

### **Adjust Fiscal Year Estimates**

The fifth step involves rebasing the weighted average application probabilities just derived to obtain a model which can be used for projection purposes. The rebasing will also be used



**Figure 1.** Comparison of weighted projections and actual application rates (from MEPCOM data) for 18 to 21 year old males from FY 1986 through 1991.



**Figure 2.** Comparison of weighted projections and actual application rates (from MEPCOM data) for 18 to 21 year old females from FY 1986 through 1991.

to include the 17 year-old group which forms a substantial proportion of accessions but could not be projected directly from YATS data. This is true because 15 year-olds are not sampled and the FY projection is roughly 2 years out from the date of the survey. Several methods of adjusting these average probabilities and the resulting numbers of applicants have been tested. Most of these methods utilized age specific probabilities along with empirical distributions of arrivals or age. These methods produced unacceptable results primarily due to the problems discussed earlier in estimating age specific application rates from the YATS respondents. For this reason a direct translation from average probabilities to annual FY rates was estimated using the same sample over which the application model was developed (1984 through 1987). Again, the 1988 and 1989 YATS are excluded and form a true out-of-sample period for model validation. This equation uses the annual FY weighted average probabilities from the cohorts to predict observed FY application rates by gender. With the exception that 17 year-old are included in the computation of the actual rates to be predicted, this amounts to regressing the values from the actual rates seen in Figures 1 and 2 on the projected rates in the figures. The resulting regression equation for males is as follows (where the coefficient t-statistics are in parentheses):

$$\begin{aligned}
 \text{Rebased rate}_{\text{males}} &= -0.0089163 + 0.9984435 \text{ Rate}_{\text{males}} \\
 &\quad (-2.080) \quad (11.628) \\
 R^2 &= 0.9854 \\
 F(1,2) &= 135.20
 \end{aligned}
 \tag{11}$$

where:

*Rebased rate*<sub>males</sub> is the final rebased projection rate for males age 17 to 21.

*Rate*<sub>males</sub> is the weighted average projection probability from Equation 10.

For females, the estimated equation for projecting age 17 to 21 application rates is:

$$\begin{aligned} \text{Rebased rate}_{\text{females}} &= 0.0020082 + 0.5259793 \text{ Rate}_{\text{females}} \\ &\quad (0.620) \quad (1.813) \end{aligned} \tag{12}$$

$$\begin{aligned} R^2 &= 0.6218 \\ F(1,2) &= 3.29 \end{aligned}$$

For females, the behavior of the 17 year-old cohort differs more from that of the 18-21 year-old cohort. For that reason a second rebasing equation for females was estimated with no attempt to expand the age range of the initial projection (the 17 year-old group is not projected). The results of that rebasing equation are:

$$\begin{aligned} \text{Rebased rate}_{\text{females 18-21}} &= 0.0009781 + 0.562812 \text{ Rate}_{\text{females}} \\ &\quad (0.500) \quad (3.210) \end{aligned} \tag{13}$$

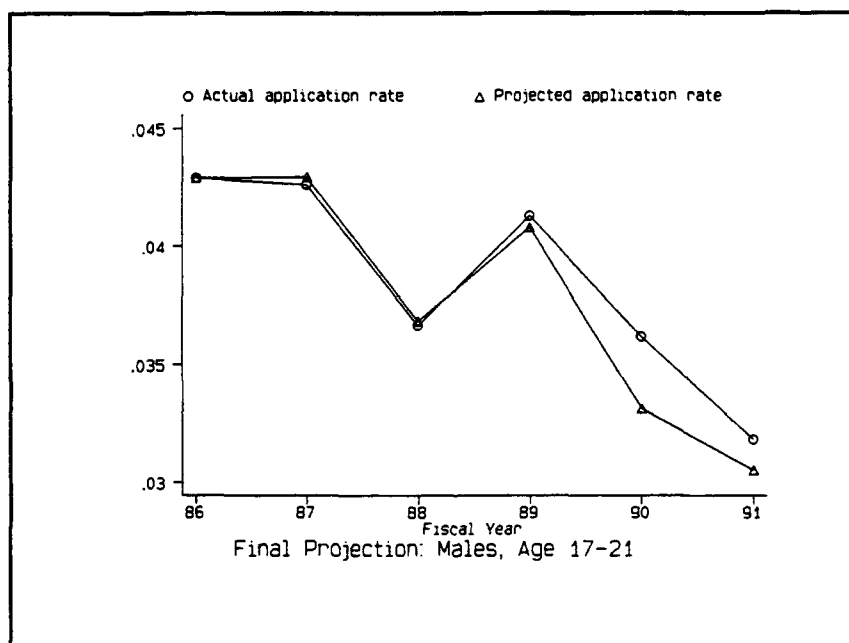
$$\begin{aligned} R^2 &= 0.8374 \\ F(1,2) &= 10.30 \end{aligned}$$

The results of applying these rebasing equations to the weighted FY estimates produced by Equation 10 are presented below in Table 40. As can be seen in the table, the largest error for males age 17 to 21 is 0.0031 (or 8.56% of the actual application rate taken from MEPCOM data). All other errors for males are less than 5.0% of the actual application rate. The Root Mean Squared Error for males over all 6 projection periods is also quite low (0.00138). The performance of the projection for males can be seen graphically in Figure 3 where the projected application rates are plotted against the actual rates from Table 40.

The performance of the full model when projecting application rates for females age 17 to 21 is also quite good, although not as strong as for males. The largest errors occur in the 1990 and 1991 out-of-sample years and are 0.00042 (5.71% of actual) and 0.00061 (10.39% of actual) respectively. Even so, an examination of Figure 4 indicates that the model captured all turning points in the application rate, including the downturn in the FY 1990 out-of-sample year.

**Table 40. Projected and Actual Annual Application Rates  
for Youth Age 17 to 21**

Fiscal Year	Male Application Rates			Female Application rates		
	Actual MEPCOM	Rebased Projection	Absolute Percentage Error	Actual MEPCOM	Rebased Projection	Absolute Percentage Error
1986	0.0429	0.0429	0.00%	0.00832	0.00848	1.92%
1987	0.0426	0.0429	0.70	0.00762	0.00789	3.54
1988	0.0366	0.0368	0.55	0.00704	0.00721	2.41
1989	0.0413	0.0408	1.21	0.00847	0.00786	7.20
1990	0.0362	0.0331	8.56	0.00736	0.00694	5.71
1991	0.0318	0.0305	4.09	0.00587	0.00648	10.39
	Root Mean Square Error (86 - 91) = .00138			Root Mean Square Error (86 - 91) = .000416		



**Figure 3. Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 17 to 21 year old males from FY 1986 through 1991.**

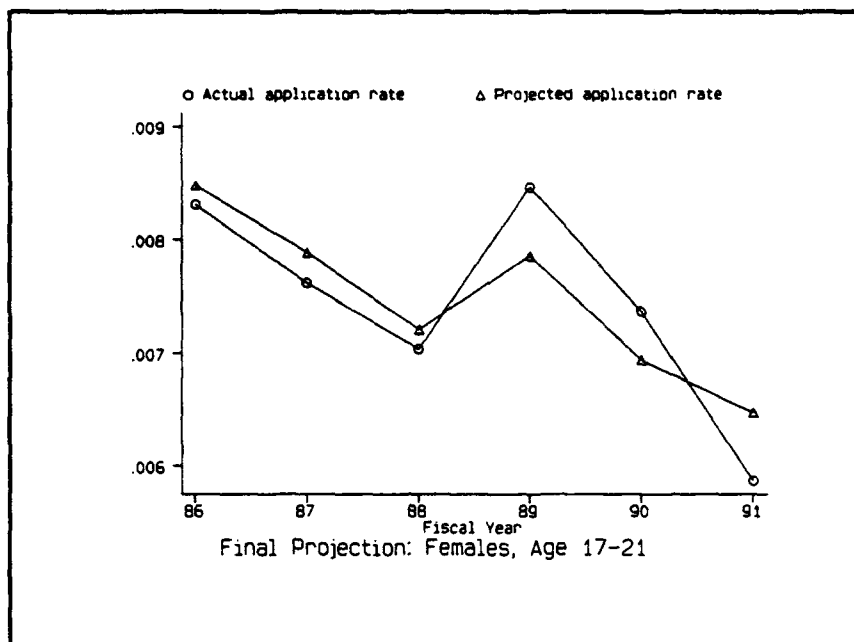
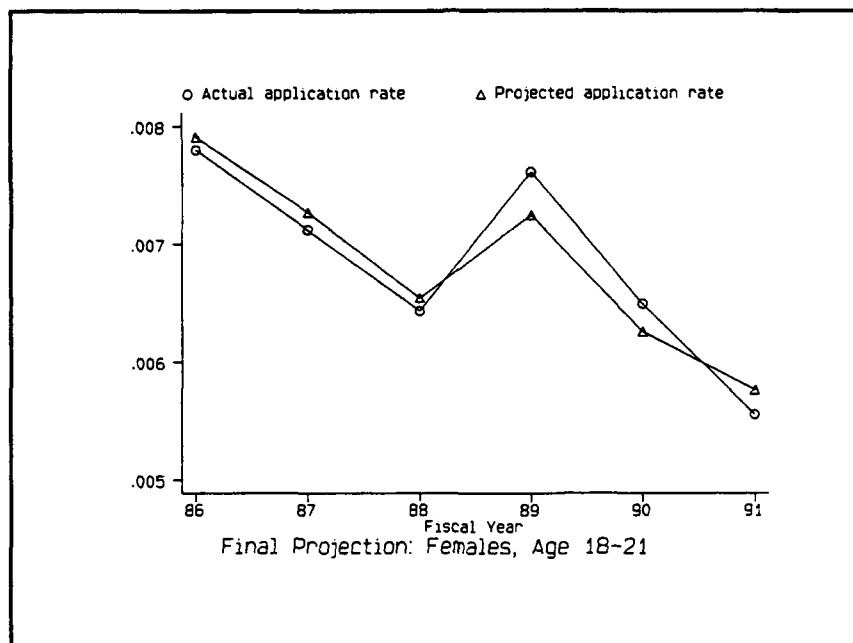


Figure 4. Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 17 to 21 year old females from FY 1986 through 1991.

Table 41. Projected and Actual Annual Application Rates for Females Age 18 to 21

Fiscal Year	Actual MEPCOM	Rebased Projection	Absolute Percentage Error
1986	.0078	.00791	1.41%
1987	.00712	.00727	2.11
1988	.00643	.00654	1.71
1989	.00761	.00724	4.86
1990	.00649	.00626	3.54
1991	.00555	.00576	3.78
Root Mean Square Error (86 - 91) = .000216			

As mentioned earlier, one factor affecting the performance of the model on females is that the 17 year age group does not follow the same temporal pattern as the estimation group (18 to 21 at application decision). For this reason, the rebasing transformation shown in Equation 13 was applied to the weighted estimates to produce a final projection for 18 to 21 year-old females (excluding 17 year-olds from the analysis). As seen in Table 41, the model performed much better when not required to extrapolate the performance of 17 year-old decision makers. No projection error is more than 5.0% of the actual application rate and both out-of-sample errors are below 4.0%. This improvement in performance can be seen in Figure 5. While the improved performance over the more constrained group helps to validate the model, it is of limited assistance in producing projections. While the 17 year-old group forms a smaller proportion of all applicants for females than it does for males, it is still a substantial component of female accessions. Despite its somewhat poorer performance, the rebasing transformation in Equation 12, which includes 17 year-old, is likely to be preferred for projection purposes.



**Figure 5.** Comparison of final, rebased projections and actual application rates (from MEPCOM data) for 18 to 21 year old females from FY 1986 through 1991.

## Methodology Summary

The steps in producing a projection from a single YATS year can be simplified somewhat when the equations for obtaining individual projections (Table 39 followed by Equation 9) and rebasing the resulting FY values (Equations 11 and 12) are taken as given. In this case, the projection of the application rate for males and females from a single YATS survey can be reduced to four steps. Table 42 summarizes the important operations in each of these steps from the preceding discussion. Completion of this process yields a projection for the first full FY following the YATS for males 17 to 21 and for females 17 to 21. Again, if the YATS survey is completed in calendar year 1991, the projection will be for FY 1993.

**Table 42. Summary of Model Methodology**

Step	Procedures
<b>Select Respondents</b>	<ol style="list-style-type: none"> <li>1) Drop all respondents without Social Security numbers.</li> <li>2) Match each remaining respondent to their earliest MEPS application record using Social Security number.</li> <li>3) Drop respondents with earliest application dates prior to their YATS interview date.</li> <li>4) Drop anyone outside the age range 16 through 19 at the date of the interview.</li> <li>5) Drop anyone who does not have valid responses to questions used to construct the variables in the logit equation from Table 39.</li> </ol>
<b>Apply Projection Equation</b>	<ol style="list-style-type: none"> <li>1) Produce linear results for each selected respondent using the coefficients from Table 39.</li> <li>2) Transform the linear result for each individual into a probability of application using the logit function from Equation 10.</li> </ol>
<b>Compute FY Estimates</b>	<ol style="list-style-type: none"> <li>1) Compute separate weighted averages of application probabilities for males and females using Equation 10. The result is two numbers — one representing the average application probability over the target FY for males and the other the average application probability for females.</li> </ol>
<b>Rebase FY Estimates</b>	<ol style="list-style-type: none"> <li>1) Rebase the average FY probabilities for males using Equation 11 to produce FY a projection of the application rate for males age 17 through 21 over the target FY.</li> <li>2) Rebase the average FY probabilities for females using Equation 12 to produce FY a projection of the application rate for females age 17 through 21 over the target FY. (Alternately, the female average probability can be rebased using Equation 13 to produce an estimate for 18 through 21 year-olds.)</li> </ol>

## VIII. CONCLUSIONS AND RECOMMENDATIONS

The development and estimation of PPMM presented in Sections III, IV, V and VI of this report provide several key conclusions. Four years of YATS data were used to estimate PPMM, 1984 through 1987, while 1988 through 1989 YATS data were used to determine the predictive capability of PPMM out-of-sample. Only four observations of time series variation were included in the data used to estimate PPMM. In addition, 1987 was the beginning of force-downsizing for DoD. Thus, DoD retention and recruiting entered a new and unique period, quite different from most of the in-sample period used to estimate PPMM. The branches of DoD were confronted with a constrained demand for recruits. Contracts declined from 349,094 in FY 1986 to 336,600 in FY 1987, followed by 307,276 in FY 1988, 303,613 in FY 1989, 265,814 in 1990, and 236,998 in FY 1991. From FY 1986 to 1991, DoD reduced contracts by 32.1%. These changes in the recruiting environment required additional analysis beyond the model estimation to enhance PPMM's predictive capability.

Several key conclusions were drawn from the development, estimation, and validation of PPMM:

1. A large percentage of the YATS respondents who were identified as applying to the military (matched a MEPS record), exhibited a date on their MEPS record prior to responding to the YATS. This phenomenon raises the question as to whether or not these respondents are representative of the intents and actions of other respondents who had not begun the application process.
2. A small percentage of the YATS respondents who were identified as applying to the military, had actually signed a contract to enter into military service prior to responding to the YATS. These respondents were excluded from the sample. The YATS instrument attempts to eliminate this type of respondent from the survey, a priori, by asking the respondent if he/she has entered into military service in the prescreening interview.
3. PPMM exhibited strong relationships with several explanatory factors. Some of the more interesting from a recruiting and accession modeling viewpoint are:
  - a. **PMIL.** PMIL exhibited a statistically significant declining contribution to the likelihood of application the less positive

the respondent's recorded propensity to enter military service.

- b. **SEMA1721.** The population of 17 to 21 year-old males (SEMA1721) exhibited a statistically significant affect for the PPMM equation for males 16 to 18 years of age. SEMA1721 indicated that low density population areas tend to exhibit relatively higher application rates.
  - c. **DFJOB.** DFJOB exhibited a negative and statistically significant effect for young and old males. The variable is based upon the individual's perception of whether he/she expects to have difficulty finding a job. This variable is a proxy for unemployment, or, at least, the respondent's perception of unemployment as it affects his/her ability to obtain employment.
  - d. **DMILCOL.** DMILCOL exhibited a statistically significant effect for young males and females. This variable identifies a part of the population which intends to go to college, but presently exhibits a positive propensity to enter the military. The implication of this relationship is that these young respondents plan to use the military as a job to accumulate funds for attending college (e.g., Army College Fund) upon separation or as a direct means to attend college while in the military.
  - e. **Q700 and Q700x2.** The grades recorded for YATS respondents were nonlinearly related to the application rate.
  - f. **Q713F and Q713Fx2.** Father's education was nonlinearly related to the application rate. The father's education recorded for YATS respondents exhibited higher tendencies for application the lower the education level of the father for young males.
  - g. **Recruiter/ASVAB contact.** Five variables were used to identify respondents who have indicated some contact or association with a recruiter and/or ASVAB. All these variables were statistically significant and positively affected the likelihood of application.
4. The actual application rates (from MEPCOM data) and the weighted projections from the PPMM Projection model track extremely close, but

the actual rates are substantially lower than the weighted projected rates. This observation is consistent with concerns about the potential for those most interested in the military self selecting into the sample at a higher rate than those who are uninterested.

5. The in-sample and out-of-sample predictive accuracy of the PPMM Projection model was very strong. In and out-of-sample projected rates and the actual application rates (from MEPCOM data) track extremely close. For 17 to 21 year-old males, projection errors ranged from 0.00% to only 8.56%. For 17 to 21 year-old females, errors ranged from 1.92% to 10.39%. The projection models for both males and females are also able to capture all turning points in the application rates, including even the downturn in the application rate in the FY 1990 out-of-sample year.
6. The multi-nomial logit results presented in Section VI were robust with explanatory variables statistically significant in many cases. The results suggest that the decision to join the military is more like the decision to go to work as compared to attending school beyond high school.

In conclusion, the PPMM Projection model presented in Section VII provides the ability to project application rates for the next fiscal year from a YATS.



## REFERENCES

## REFERENCES

- Becker, G.S. (1971). *Economic Theory*. Alfred A. Knopf, New York.
- Cotterman, R.F. (1986). *Forecasting enlistment supply: a time series of cross sections model*. The RAND Corporation, R-3252-FMP.
- Curtis, E.W., Borack, J.I., and Wax, S.R., (1987). *Estimating the Youth Population Qualified for Military Service*. Navy personnel Research and Development Center, NPRDC TR 87-32. San Diego, California.
- Dale, D. and Gilroy, C. (1984). Determinants of enlistments: a macroeconomic time-series view. *Armed Forces & Society*, Vol.10, No.2.
- Daula, T.V., Fagan, T.W., and Smith, D.A., (1982). *A Microdata Model of Enlistment in the Armed Forces*. Paper presented at the summer meeting of the Econometric Society, Ithaca, New York.
- DeVany, A.S., & Saving, T.R., (1982). Life-cycle job choice and the demand and supply of entry level jobs: some evidence from the Air Force. *Review of Economics and Statistics*, LXIV(3).
- Fast, J.C., Stone, B.M., and Looper, L.T., (1989). *Civilian Availability and Manpower Requirements System*. Final Report, Contract No. F33615-83-C-0030, Task No. 0030-30. Brooks Air Force Base, Texas: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Goldberg, L. (1988). *Analysis of military enlistments in the 1980s*. Final Report, Contract No. MDA 903-87-C-0790. Office of the Assistant Secretary of Defense, (FM&P/MM&PP) Accession Policy, Pentagon, Washington, D.C.
- Hosek, J.R., and Peterson C.E. (1985). *Enlistment Decisions of Young Men*. The RAND Corporation, R-3238-MIL.
- House, D.R., Saving, T.R., and Stone, B.M. (1985a). *Employment of Recent Dental Graduates: Phase II*. Department of Health and Human Services, Washington, D.C.
- House, D.R., Saving, T.R., and Stone, B.M. (1985b). *Impact of Dental Prepayment on the Demand for Dental Care*. Subcontract with Research Triangle Institute for the Department of Health and Human Services.
- Orvis, B. R., (1982). *Forecasting Enlistment Actions from Intention Information: Validity and Improvement*. The Rand Corporation, N-1945-MRAL.

- Orvis, B. R., (1984). *Analysis of Youth Cohort Enlistment Intention Data*. The Rand Corporation, N-2076-MIL.
- Orvis, B. R., (1986). *Relationship of Enlistment Intentions to Enlistment in Active Duty Services*. The Rand Corporation, N-2411-FMP.
- Orvis, B.R. and Gahart, M.T., (1985). *Relationship of Enlistment Intention and Market Survey Information to Enlistment in Active Duty Military*. The RAND Corporation, N-2292-MIL.
- Orvis, B.R. and Gahart, M.T., (1989). *Quality-Based Analysis Capability for National Youth Surveys: Development, Application, and Implications for Policy*. The RAND Corporation, R-3675-FMP.
- Orvis, B.R., Gahart M.T., and Hosek, J.R. (1989). *Predicting Enlistment for Recruiting Market Segments*. The RAND Corporation, N-2852-FMP.
- Reed, W.R. and Saving, T.R. (1987). *Compensating Wage Differentials and Retention. Proceedings: USAF/RAND Entitlements Planning Conference May 28-29, November 6-7, 1986*. The RAND Corporation.
- Saving, T.R., Battalio, R.C., DeVany, A.S., Dwyer, G.P., and Kagel, J.K., (1980). *Air Force Enlisted Personnel Retention-Accession Model*. (AFHRL-TR-80-12). Brooks Air Force Base, Texas: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Saving, T.R. and Stone, B.M., (1983). *Air Force Prior Service Market Analysis*. AFMPXOA, Pentagon, Washington, DC.
- Saving, T.R., Stone, B.M., Looper, L.T., and Taylor, J.N., (1985). *Retention of Air Force Enlisted Personnel; An Empirical Examination*. (AFHRL-TP-85-6). Brooks Air Force Base, Texas: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Stone, B.M., Saving, T.R., Turner, K.L., and Looper, L.T., (1990). *Integrated Economic and Behavioral Modeling of Accession and Retention*. Brooks Air Force Base, Texas: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Theil, H. (1971). *Principles of Econometrics*. Amsterdam: North-Holland.
- U.S. Department of Education (1986). *High school graduation rates*. National Center for Education Statistics. Washington, D.C.

Verdugo, N., and Berliant, K.R., (1988). *Estimating the Army's prime recruiting market*. U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginia.

**APPENDIX A. REVIEW OF ANALYSIS OF INTENTION  
INFORMATION**

## REVIEW OF ANALYSIS OF INTENTION INFORMATION

### *Forecasting Enlistment Actions from Intention Information: Validity and Improvement*

Bruce R. Orvis, December 1982

#### **Purpose**

To investigate the relationship between survey enlistment intention measures and respondents' subsequent enlistment decisions, to assess the usefulness of including intention variables in enlistment decisions models based on demographic and economic factors, and to provide guidance in designing intention measures for the particular enlistment issue of interest.

In this phase the research was primarily focused on linking enlistment decision information with existing enlistment intention data. The results from this analysis were used to quantify the relationship between enlistment intentions and actual decisions. Several intention measures were examined to distinguish measures of the propensity to enlist in the military in general as opposed to the individual services. The next major effort was to focus on the enlistment intention results. The Applicant and YATS surveys both ask questions pertaining to how likely is it that the respondent will be serving in the military in the next few years. Also, the YATS survey contains service specific intention questions. In addition, the YATS survey asks an unaided mention question of what the respondent plans to do in the next few years.

#### **Data**

The survey data were drawn from the 1981 Applicant Survey of males taking the written test to enter the military and from 10 semi-annual waves of the Youth Attitude Tracking Survey (YATS) of males ages 16-21, covering the period Spring 1976 to Fall 1980. Enlistment data were obtained from Defense Manpower Data Center (DMDC) extracts of the AFEES reporting system records maintained by the Military Enlistment Processing Command (MEPCOM). These data were used as follow-ups to determine actual enlistment decisions. The Applicant Survey and each YATS wave matched approximately 3500 respondents.

## **Summary**

In this study, the analysis focused on examining the characteristics of enlistment applicants and their distributions over time, and on investigation of the validity of intention measures on high and low quality recruits, short and long-term decisions. The results of this study are based on the first five waves of the YATS survey.

To examine these issues, distributions of the respondent's enlistment actions over time were examined. These distributions were compared to the distribution of those taking the written test to enter military service along with the number of actual enlistments. The data show that the number of enlistments and written exams increased throughout the 42 month period. This initial work indicates the need for long-term follow-ups of studies modeling enlistment decisions. It also suggests that recruiters could benefit from long-term follow-ups as well.

In addition, the characteristics of the respondents who took the written test were compared to those of respondents in the population as a whole. In general, the persons taking the test tended to be younger, were less likely to be high school graduates, and were more likely to be from a minority background. Overall, the data suggest that enlistment intention measures are valid for both high and low quality applicants and retain their most discriminatory power in the first twelve months. Finally, the service specific measures were shown to perform better than the general measure.

## **Applicant Survey Results**

There were five responses coded in the intention question. They were "definitely will," "probably will," "probably will not," "definitely will not" and "don't know." The first two categories were coded as the positive propensity group and the remaining three groups as negative propensity. These respondents were further stratified into high and low quality applicants. High quality applicants are high school diploma graduates who score in the upper 50th percentile of the written test, categories I-IIIa. All others are coded as low quality applicants. Actual enlistments were compared to the intention measure. For the high quality applicants, 53% of the 'definitely will serve' and 27% of the 'probably will serve' applicants enlisted in the military one year later. Approximately 7% of the negative propensity group enlisted. Even though the absolute levels were lower for low quality applicants, the pattern was

similar for low quality applicants. Once the applicants who did not qualify for enlistment (i.e. did not pass the written or physical test) were accounted for, the patterns of enlistment rates for both the high and low quality groups did not differ statistically. In general, a higher percentage of the 'definitively will serve' category of applicants enlisted for military duty. This was followed by the 'probably will' and the negative propensity groups. Overall, 52% of the high and low quality applicants qualified to serve enlisted within the one year follow-up period. The enlistment rates for the 'probably will serve' group, and ranged from 26-21% and ranged from 7-11% for the negative propensity group.

### **YATS Results**

The above analysis was conducted on applicants who were generally well along their way in the application process. All these applicants had taken the written exam and many had taken the physical exam as well. Using YATS data, they now proposed to examine a national sample of youth whose enlistment decision was thought to be years away. Data from the first five waves (Spring 1976 to Fall 1978) of the YATS were used in order to provide a reasonably long follow-up period that extended through December 1981. Initial results show that about 3% of the sample said they definitely will serve, 24% probably will serve, and 73% expressed negative propensity. Of these groups, 33, 17, and 5% respectively did enlist by 1981. Next, a measure that looked at the enlistment propensity of the individual services was analyzed. The results were similar to the general measure with the exception that more people were classified in the positive propensity group.

In order to better discriminate among the different propensity groups, the results of the unaided mention question along with those of the general measure were combined to form a composite measure for the four groups. The four groups are 1) unaided mention and definitively will serve, 2) unaided mention and probably will serve, 3) no unaided mention and probably will serve, and 4) no unaided mention and definitively or probably will not serve. Looking at the enlistment results they found that 49, 32, 15, and 5% of the persons in groups 1-4 respectively enlisted within a year. Also for the positive groups, the number of persons taking the written test experienced a similar pattern with a higher overall magnitude. In general, the composite measure was shown to distinguish persons with different enlistment rates for the 42 month

follow-up period. Further analysis was conducted on the individual services. The comparison was between unaided mention and propensity to serve in the military versus the propensity to serve in an individual service. The results were very similar and indicate that the service specific measures perform better than the general measure in terms of predicting enlistment behavior. As an example, 33% of the individuals with an unaided mention and definite intention to serve in the Army actually enlisted one year later. This compares to 18% enlistment rate for persons with an unaided mention and definitive intention to serve in the military. The analogous comparison for the probable intention group yields a comparison of 19 to 13%. The results were similar across services.

### ***Analysis of Youth Cohort Enlistment Intention Data***

Bruce R. Orvis, June 1984

#### **Purpose**

The purpose was to investigate the relationship between survey enlistment intention measures and respondents' subsequent enlistment decisions, to assess the usefulness of including intention variables in individual and aggregate enlistment decisions models based on demographic and economic factors, to include both enlistment and individual models of first-term performance, and to develop methods of identifying the quality of survey respondents who have not taken the ASVAB.

#### **Data**

The survey data were drawn from 12 semi-annual waves of the YATS, covering the period Spring 1976 to Fall 1982. Additional data were collected from randomly selected nonprior service males ages 16-21. These data were combined with YATS data from a YATS study. Enlistment data were obtained from the Defense Manpower Data Center (DMDC) extracts of the AFEES reporting System records maintained by the Military Enlistment Processing Command (MEPCOM). 40,993 records were matched yielding 3102 enlistees through December 1982. Aggregate level intention data were merged for the fifteen largest states over a 58 month period. The composite intention measure in the previous study was used.

## Summary

This report was meant to build upon previous work. Even though a strong relationship exists between enlistments and intentions, it is possible that intention information is completely captured by other demographic factors. If this were the case, then intention information will not add to the predictive capability of the enlistment behavior model. To test this proposition, data from the first seven waves of the YATS survey, followed up through the end of 1982 were used. A variety of demographic factors such as age, education, father's education, AFQT score, DEP and work history, math and science courses taken, race, geographic location, and others were combined with the intention data.

In order to analyze enlistment rates across the four groups, regression analysis was performed on each group, and comparisons were made with the negative propensity group. After controlling for a variety of demographic factors, the enlistment rate for the unaided mention-definite intention group increased 36% over the negative propensity group. Similarly, the increase in the enlistment rate was 23 and 7% for the unaided mention-positive propensity, and the positive propensity-no unaided mention. A corresponding analysis was conducted that examined the percent of respondents taking the ASVAB. The resulting increase in percentage was 41, 33, and 12% respectively. Thus, intention data were shown to add to the predictive capability of the enlistment model that includes demographic characteristics.

Given that intention data add to the ability to predict enlistment decisions of YATS respondents, the next step was to assess whether or not it could help in predicting first-term performance. The performance measures studied were attrition and promotion. These measures were regressed on the same demographic factors. No significant relationship was found to exist between enlistment intentions and promotion. The results showed that length of time in service was the strongest contributing factor in promotions. However, an examination of attrition showed that attrition was significantly lower for the highest propensity group. The results also showed no significant difference between the attrition rates of the negative propensity and the positive propensity-no aided mention groups, the largest two cohorts. In addition, attrition rates were shown to vary by high school status.

The final contribution of the paper focuses on the ability of the model to predict AFQT scores based on demographic characteristics. Similar to previous work, YATS respondents'

AFQT percentile scores were found to have a positive relationship with grade point average, the number of math courses completed, and father's education. AFQT scores were found to be lower for minorities and respondents from the South. The model was able to correctly classify about 70 to 75% of the respondents into either the upper or lower 50th percentile. The same analysis was performed using the quality index score to predict the AFQT score. This model was only able to accurately classify about 65% of the respondents. Whereas the AFQT measure correctly classified 74% in the upper percentile group, the quality index classified only 50%. Finally, a model was constructed to determine if intention data in conjunction with aggregate data could better explain regional enlistment rates. The results suggest that intention information improves aggregate forecasts across regions.

### ***Relationship Of Enlistment Intention and Market Survey Information to Enlistment in Active Duty Military Service***

Bruce R. Orvis and Martin T. Gahart, June 1985

#### **Purpose**

The purpose was to investigate the relationship between survey enlistment intention measures and non prior service male respondents' subsequent enlistment decisions, to develop methods of identifying the quality of survey respondents who have not taken the ASVAB, and to investigate the relationship between aggregate intention levels and aggregate enlistment rates.

#### **Data**

The survey data were drawn from 13 semi-annual waves of the YATS, covering the period Spring 1976 to Fall 1983. The data were collected from 5250 randomly selected nonprior service males (NPS) ages 16-21. Enlistment data were obtained from the Defense Manpower Data Center (DMDC) extracts of the AFEES reporting System records maintained by the Military Enlistment Processing Command (MEPCOM). In addition, 3700 records of NPS males ages 16-21 from the 1979-1982 waves of the National Longitudinal Surveys of Youth Labor Market Experiences were analyzed. The remaining data base used in this report was developed by Robert Cotterman, formerly of Rand. This database groups the 50 states into 17 regions and covers the October 1976-March 1983 period.

## Summary

The results from the unaided mention question and the general intention measure used in previous work were combined to form a composite measure with three categories. The first category combined the unaided mention with definitely or probably will serve; the second group had no unaided mention and definitely or probably will serve; and the remaining category was composed of persons with negative intent or don't know. Preliminary evidence from the Spring 1976-Fall 1980 YATS showed that 37, 15, and 6% of the respondents from groups 1-3 respectively had enlisted in the military after a minimum follow-up of 42 months. The results with respect to the number testing exhibited the same pattern.

Next, regression analysis was used to explain differences in enlistment and testing rates across cohorts. The enlistment and testing rates in the positive intention groups were compared to those in the negative intention group. After controlling for differences in background characteristics, the analysis revealed that persons in group 1, on average, had enlistment rates 24% higher and testing rates 30% higher than the control group. Group 2 had rates 5 and 9% higher respectively. Further analysis revealed that even though only 6% of the individuals with negative intentions enlisted in military service, they accounted for 46% of total enlistees. This result is mainly attributed to the size of this group. This fact indicates that small increases in the enlistment rate for this group can account for a substantial increase in the number of enlistees. Thus, the characteristics of enlistees in each group need to be analyzed.

Analysis within both the positive and negative intention groups for persons in high school was performed on YATS data to address the issue of whether or not the model could distinguish enlistees from non-enlistees. Both long-term and short-term factors were identified to influence the enlistment decision in the two groups. In many cases the factors identified were the same. In general, those individuals who were minority status, had taken fewer math courses, had difficulty finding work, discussed enlisting in military service, and perceived the military as offering job security had higher enlistment rates. This analysis was then performed on non high school respondents. As before, the results with respect to the negative and positive intention groups was similar. Additional analysis was performed on juniors using NLS data. As shown by YATS, among respondents with negative intentions, those not on a college track were more likely to enlist. Individuals attending vocational schools, those having difficulty finding a job,

and younger respondents were also shown to have higher enlistment rates. Those individuals making more money and who liked their jobs were less likely to enlist.

The next contribution of this report examines the relationship between aggregate enlistment rates and aggregate intention levels in current and future periods. Using the YATS and Cotterman data bases, a time series cross sectional analysis was performed for 17 geographical regions. These results indicate that there is a significant relationship between intention levels and concurrent high quality enlistments. Next, analysis was performed to determine if there were any lagged effects of intention on enlistments. After controlling for current intention, the lagged intention measure was significant in three out of four analyses.

Finally, work was updated on identifying the quality of YATS respondents who had not taken the ASVAB. Based on background characteristics, a model was developed that correctly classified between 70-75% of the respondents in the appropriate AFQT category. These results compared favorably to the more complex methods of predicting AFQT scores. In addition, the results indicate that this simple measure is reliable over time.

### ***Relationship of Enlistment Intentions to Enlistment in Active Duty Services***

Bruce R. Orvis, September 1986

#### **Purpose**

Since women were not included in YATS surveys until the Fall 1980, initial work by Orvis had examined the relationship between survey enlistment intention measures and non prior service male respondents' subsequent enlistment decisions. This report is intended to highlight the results of the analysis of men and to provide the same analysis of women.

#### **Data**

The data were drawn from 11 waves of the YATS survey administered between Spring 1976 and Fall 1981. This file contained records on 37,047 males and 6,226 females. Enlistment data were obtained from the Defense Manpower Data Center (DMDC) extracts of the AFEES reporting System records maintained by the Military Enlistment Processing Command (MEPCOM). The follow-ups extended through March 1985.

## Summary

Using the results from the Fall 1980 and Fall 1981 survey waves, the enlistment/intention analysis was performed. The measures used were the same intention and unaided mention as used in previous studies. The data show that approximately 87% of the women expressed negative intent, and about 13% expressed positive intent to join the military. This compares to about 68 and 32% respectively for men in the same period. An examination of the distributions of the number of enlistees across the three groups, 1) positive intention and unaided mention 2) positive intention and no unaided mention and 3) negative intention; reveal that 18, 4, and 1% of the women in groups 1-3 respectively enlisted by the follow-up period. This compares to 31, 12 and 6% respectively for men. The same pattern for both men and women was found to be similar with respect to the percent of respondents taking the ASVAB. It is hypothesized that demand constraints on the types of jobs available, physical requirements, and social pressures tend to limit enlistment among women. Further analysis revealed, after controlling for differences in background characteristics, that the influence of positive intentions on the testing rate was similar for both sexes. The results with respect to enlistments were also similar though not as strong as those with respect to intentions.

Overall, the data revealed that despite the low enlistment rate for women with negative intentions, this group accounted for 63% of actual enlistees. Fourteen percent came from the positive intention and unaided mention group and the remaining 23% from the positive intention and no unaided mention. This compared to 46, 19, and 35% respectively for men. The implication of this analysis revealed that when we combine intention data for both sexes, we are likely to overstate women's interest in enlisting.

## ***Predicting Enlistment for Recruiting Market Segments***

Bruce R. Orvis, Martin T. Gahart, and Hosek, P.R., September 1989

### **Purpose**

The purpose of this report is to synthesize the results of previous work and to examine the question of whether geodemographic clusters add to the predictive power of models of individual enlistment decision making. In addition, this report investigates whether the factors used in predicting geodemographic models vary across geodemographic groups. Finally, this report examines the relationship between ACORN (A Clustering of Residential Neighborhoods) information and micro models of enlistment behavior.

### **Data**

The data were drawn from 11 waves of the YATS survey administered between Spring 1976 and Fall 1981. The file contained records on 36,648 individuals. Enlistment data were obtained from the Military Entrance Processing Station (MEPS) Reporting Systems. The follow-ups extended through March 1985. A second, choice-based sample of youth ages 16-22 was constructed from the 1979 Air Force Entrance and Examining Station (AFEES) and NLS Surveys. This file contains 4443 records of enlistees from the AFEES and 1093 records of nonenlistees from the NLS files. The third major database consisted of ACORN cluster information that reflects the percentage of male youth population in each of the 44 ACORN clusters for each zip code and FIPS code in the U.S.

### **Summary**

Since research up to this point had only validated the use of ZIP-level information, this paper begins by examining the relationship between estimates of ACORN information produced by the FIPS and ZIP code approaches. The results of this analysis showed that the FIPS and ZIP code measures produce similar results. Having determined the above, the next effort was to determine the relationship between geodemographic information such as that contained in ACORN and information from individual-level micro models such as that contained in YATS-MRS or AFEES-NLS.

Using the YATS-MRS micro model, comparisons between actual and predicted enlistment rates for the 44 ACORN clusters within each county were made. Results of the analysis show that there is a highly significant relationship between observed and predicted enlistment rates using the product-moment and the rank correlation measures. The correlations were 0.84 and 0.81, respectively. Thus, the YATS micro models accounted for most of the variation in enlistment rates among the ACORN clusters. Likewise, analysis was performed using the AFEES-NLS data base. This comparison produced similar results as the correlation measures were estimated at 0.73 and 0.76, respectively. This analysis confirms the previous finding that the variables in the micro model do a good job of explaining the variation in enlistment rates across geodemographic clusters.

Finally, the data were used to determine if ACORN information could improve the predictive capability of the micro models. Stepwise regression was performed that provided little evidence of a significant increase in explanatory power of the geodemographic information. In addition, three logistic regression equations were analyzed that used ACORN information, micro-model information, and both ACORN and micro-model information. The predicted probabilities from each model were ranked and grouped into five quintiles. The results of this analysis showed that even though ACORN information alone accounted for some variation in enlistment rates, the micro-model accounted for much more of the variation. Also, when the ACORN information was added to the micro-model, there was no statistically different prediction in the enlistment rates across the five quintiles.

## **APPENDIX B. RE-ESTIMATION OF THE QUALITY EQUATION**

## RE-ESTIMATION OF THE QUALITY EQUATION

### Orvis and Gahart, 1989

Orvis and Gahart (1989) built upon the work initiated by Orvis (1984). The primary purpose of Orvis and Gahart's study (1989) was to develop a methodology to estimate the probability that respondents would score in the upper 50th percentile taking the Armed Forces Qualifying Test (AFQT). It also provided a method of identifying and estimating the number of high-quality youths from the surveyed population. Two sets of equations were estimated to identify high and low aptitude applicants. The first set of equations estimated the probability that a given individual would take the AFQT. The second set of equations estimated the probability, given that an individual took the AFQT, that he would score in the upper 50th percentile.

The results for the high school testing equation for males summarized below are consistent with Orvis and Gahart (1985). In general, the coefficients on background and economic factors indicate that blacks are more likely to test as opposed to white non-Hispanics. Individuals with positive expectations about job searches are less likely to test. In addition, the testing rate was negatively correlated to grade-point average (GPA) and there was no difference in regional variations. Also, the likelihood of testing was directly related to positive intentions and prior contact with a recruiter. White non-Hispanic, non Southern resident, and level of father's education, were found to be positively related to scoring in the upper 50th percentile. Also, GPA, the number of math courses completed, and high school senior status were found to be positively related to scoring well. However, even though positive intentions were positively related to the testing rate, they were found to be negatively related to scoring in the upper 50th percentile. The results for the respondents not in high school were similar to those in high school. Similar analysis was attempted for the female respondents. However, given the small sample size for females it proved impossible to estimate the model as most of the variables were insignificant.

Further analysis was conducted to assess the accuracy and reliability of the estimation procedure. First, analysis was conducted that compared the actual proportion of AFQT test

takers scoring in the upper 50th percentile with the predicted probability of doing so. Secondly, analysis compared the actual to predicted characteristics of persons scoring in the upper 50th percentile. The results showed that actual distributions of the probabilities of taking the AFQT and the characteristics of the test takers were in close agreement with their predicted probabilities and characteristics.

### **Re-estimation of the Quality Equation**

To replicate the work of quality estimation of Orvis and Gahart with respondents to later YATS requires some modification in the specification of the equation since some of the variables originally used by Orvis are no longer collected in the YATS surveys. Where possible, proxies were used. The quality equation was re-estimated using male respondents, age 16 to 21, from the 1984 to 1987 YATS. Out-of-sample predictions were done using male respondents from the 1988 and 1989 YATS. The equation was estimated using a logit function.

The dependent variable for the estimation will be a binary zero/one variable. The dependent variable will be equal to one if the respondent scored above the 50th percentile, and zero if the respondent scored below the 50th percentile. Table B1 details the explanatory variables used in the re-estimation of the quality equation. All of the variables in Table B1 are binary with the exception of *Q700* (grades). Results of the estimation are provided in Table B2 and are similar to the results of Orvis and Gahart in expected relationships.

Many of the explanatory variables in the equation were significant at the 99% level of significance. As with the Orvis and Gahart study, black and other racial groups were found to be less likely to be in the upper 50th percentile based on AFQT scores. Respondents with a high school diploma were more likely to be in the upper 50th percentile. No effect was observed from the regional binary variables (statistically insignificant). The respondent's father having a high school diploma or further education increased the likelihood of the respondent being in the upper 50th percentile. Respondents' grades and courses completed in school also affected the likelihood of being in the upper 50th percentile. The equation predicted well in and out-of-sample. The equation predicted correctly 73% of the time in-sample. Out-of-sample the equation predicted correctly 77% of the time in 1988 and 72% of the time in 1989.

**Table B1. Variable Definitions for Quality Equation**

Variable Name	Description	YATS Questions
<b>AFQTG1</b>	Binary: 1 = Score in the 50th percentile or greater on AFQT, 0 = Otherwise	AFQT Percentile from matched MEPS record
<b>DBLACK</b>	Binary: 1 = Black, 0 = otherwise	Q714 is "2" for black
<b>DOTHER</b>	Binary: 1 = Non-Caucasian or non-black, 0 = Otherwise	Q714 is equal to "3" for Asian or Pacific Islander or "4" for American Indian or Alaskan Native and if Q714 equals "1" for Caucasian and Q715 equals "1" for Hispanic background
<b>DDIPLOMA</b>	Binary: 1 = High school diploma, 0 = Otherwise	Q406 is equal to "1" which is a high school diploma
<b>DSOUTH</b>	Binary: 1 = Census district South, 0 = Otherwise	STFIPS
<b>DNORTH</b>	Binary: 1 = Census district North Central, 0 = Otherwise	STFIPS
<b>DWEST</b>	Binary: 1 = Census district West, 0 = Otherwise	STFIPS
<b>DFEDUC</b>	Binary: 1 = Father has greater than a high school diploma, 0 = Otherwise	Q713F is greater than "12"
<b>DFHS</b>	Binary: 1 = Father has a high school diploma, 0 = otherwise	Q713F is equal to "12"
<b>Q700</b>	Grades: Values 1 to 7	Q700 is "1" for mostly A's, "2" for mostly A's and B's, "3" for mostly B's, "4" for mostly B's and C's, "5" for mostly C's, "6" for mostly C's and D's, and "7" for mostly D's and F's
<b>DQ702</b>	Binary: 1 = Taken elementary algebra, 0 = Otherwise	Q702 is equal to "1"
<b>DQ703</b>	Binary: 1 = Taken plane geometry, 0 = Otherwise	Q703 is equal to "1"
<b>DQ706</b>	Binary: 1 = Taken intermediate algebra, 0 = Otherwise	Q706 is equal to "1"
<b>DQ707</b>	Binary: 1 = Taken trigonometry, 0 = Otherwise	Q707 is equal to "1"

**Table B2. Quality Equation Re-estimation Results**

Number of Observations	2,018
Log Likelihood	-1095.53
Chi-Square	604.28

<u>Name</u>	<u>Coefficient</u>	<u>t-statistic</u>	<u>Prob &gt;  t </u>	<u>Mean</u>
AFQTG1				0.5164729
DBLACK	-0.437327	-10.930	0.000	0.1639262
DOTHER	-0.286105	-3.910	0.000	0.0329727
DDIPLOMA	0.121638	4.501	0.000	0.5870841
DSOUTH	-0.029523	-0.887	0.375	0.3574675
DNORTHC	-0.046399	-1.160	0.246	0.1732905
DWEST	-0.009594	-0.242	0.809	0.1849441
DFEDUC	0.209466	5.635	0.000	0.3277252
DFHS	0.114133	3.315	0.001	0.4561648
Q700	-0.051887	-4.629	0.000	3.7222390
DQ702	0.191496	5.247	0.000	0.7916930
DQ703	0.194317	6.575	0.000	0.5589466
DQ706	0.077499	2.588	0.010	0.4976285
DQ707	0.198971	5.248	0.000	0.2190455
CONSTANT	-0.222437	-3.371	0.001	1.0000000



## **APPENDIX C. CREATION OF VARIABLES FROM YATS DATA**

## CREATION OF VARIABLES FROM YATS DATA

### Replication of the Orvis Equation

Variable Name	Variable Description	YATS Questions
<b>PMIL1</b>	Binary: 1 = Unaided mention, definite intention, 0 = Otherwise	Q438 is military and Q503 is "definitely"
<b>PMIL2</b>	Binary: 1 = Unaided mention, probable intention, 0 = Otherwise	Q438 is military and Q503 is "probably"
<b>PMIL3</b>	Binary: 1 = No unaided mention, definite or probable intention, 0 = Otherwise	Q438 is not military and Q503 is "definitely" or "probably"
<b>Q403</b>	Age: Values 16 to 21	Q403
<b>DSOUTH</b>	Binary: 1 = Census district South, 0 = Otherwise	STFIPS2
<b>DNORTHC</b>	Binary: 1 = Census district North Central, 0 = Otherwise	STFIPS2
<b>DWEST</b>	Binary: 1 = Census district West, 0 = Otherwise	STFIPS2
<b>DBLACK</b>	Binary: 1 = black, 0 = otherwise	Q714 is "2" for black
<b>DOTHER</b>	Binary: 1 = non-Caucasian or non-black, 0 = otherwise	Q714 is equal to "3" for Asian or Pacific Islander or "4" for American Indian or Alaskan Native and if Q714 equals "1" for Caucasian and Q715 equals "1" for Hispanic background
<b>DDIPLOMA</b>	Binary: 1 = high school diploma, 0 = Otherwise	Q406 is equal to "1" which is a high school diploma
<b>DISCHOOL</b>	Binary: 1 = still attending school, 0 = Otherwise	Q407 is equal to "1" if the respondent is currently attending school or will be in the fall
<b>DQ702</b>	Binary: 1 = Taken elementary algebra, 0 = Otherwise	Q702 is equal to "1"
<b>DQ703</b>	Binary: 1 = Taken plane geometry, 0 = Otherwise	Q703 is equal to "1"
<b>DQ705</b>	Binary: 1 = Taken computer science, 0 = Otherwise	Q705 is equal to "1"
<b>DQ706</b>	Binary: 1 = Taken intermediate algebra, 0 = Otherwise	Q706 is equal to "1"

<b>DQ707</b>	Binary: 1 = Taken trigonometry, 0 = Otherwise	Q707 is equal to "1"
<b>DQ708</b>	Binary: 1 = Taken calculus, 0 = Otherwise	Q708 is equal to "1"
<b>DQ709</b>	Binary: 1 = Taken physics, 0 = Otherwise	Q709 is equal to "1"
<b>Q700</b>	Grades: Values 1 to 7	Q700 is "1" for mostly A's, "2" for mostly A's and B's, "3" for mostly B's, "4" for mostly B's and C's, "5" for mostly C's, "6" for mostly C's and D's, and "7" for mostly D's and F's
<b>Q713F</b>	Father's education: Values 7 to 20	Q713F is "7" for less than 8th grade, "8" for 8th grade, "9" for 9th grade, "10" for 10th grade, "11" for 11th grade, "12" for 12th grade, "13" for 1st year college/junior college or community college/vocational, business, or trade school (Freshman), "14" for 2nd year college/junior or community college/vocational, business, or trade school (sophomore), "15" for 3rd year of 4-year college (junior), "16" for 4th year of 4-year college (senior), "17" for 5th year college/1st year graduate or professional school, "18" for 2nd year graduate or professional school, "19" for 3rd year graduate or professional school, "20" more than 3 years graduate or professional school.
<b>DCWORK</b>	Binary: 1 = currently working, 0 = Otherwise	Q416 is "1" if employed full or part-time
<b>DLWORK</b>	Binary: 1 = currently not employed and looking for work, 0 = Otherwise	Q417 is "1" is looking for work and Q416 is "2" (currently not employed)
<b>DFJOB</b>	Binary: 1 = no difficulty finding a job, 0 = Otherwise	Q436 is "4" if it is not difficult to find a full time job in the community
<b>DRECR</b>	Binary: 1 = have contacted a recruiter, 0 = Otherwise	Q628 is "1" if ever talked to a recruiter

### Extension of the Orvis Intention to Apply Variables

Variable Name	Variable Description	YATS Questions
<b>PMIL3</b>	Binary: 1 = No unaided mention, definite intention, 0 = Otherwise	Q438 is not military and Q503 is "definitely"

<b>PMIL4</b>	Binary: 1 = No unaided mention, probable intention, 0 = Otherwise	Q438 is not military and Q503 is "probably"
<b>PMIL5</b>	Binary: 1 = No unaided mention, probably not intention, 0 = Otherwise	Q438 is not military and Q503 is "probably not"

### Including Males and Females in the Application Equation

Variable Name	Variable Description	YATS Questions
<b>DMALE</b>	Binary: 1 = Male, 0 = Female	Q402 is 1 for "male" or 2 for "female"

### New Intention Variables for the Estimation of PPM

Variable Name	Variable Description	YATS Questions
<b>PMIL41</b>	Binary: 1 = Unaided mention, definite or probable intention, and definite or probable intention for at least one of the four services, or guard or reserve component, 0 = Otherwise	Q438 is military, Q503 is "1" (definitely) or "2" (probably), and highest response of CPYATS82, Q505, Q507, and Q509 is "1" (definitely) or "2" (probably)
<b>PMIL42</b>	Binary: 1 = No unaided mention, definite or probable intention, and definite or probable intention for at least one of the four services, or guard or reserve component, 0 = Otherwise	Q438 is no military, Q503 is "1" (definitely) or "2" (probably), and highest response of CPYATS82, Q505, Q507, and Q509 is "1" (definitely) or "2" (probably)
<b>PMIL43</b>	Binary: 1 = No unaided mention, probable or probably not intention, and probable, probably not, or definitely not intention for at least one of the four services, or guard or reserve component, 0 = Otherwise	Q438 is no military, Q503 is "2" (probably) or "3" (probably not), and highest response of CPYATS82, Q505, Q507, and Q509 is "1" (definitely), "2" (probably) or "3" (probably not); or Q503 is "4" (definitely not), and highest response of CPYATS82, Q505, Q507, and Q509 is "1" (definitely) or "2" (probably)
<b>PMIL44</b>	Binary: 1 = No unaided mention, definitely not intention, and probably not or definitely not intention for at least one of the four services, or guard or reserve component, 0 = Otherwise	Q438 is no military, Q503 is "3" (probably not) and highest response of CPYATS82, Q505, Q507, and Q509 is "4" (definitely not); or Q503 is "4" (definitely not), and highest response of CPYATS82, Q505, Q507 and Q509 is "3" (probably not) or "4" (definitely not)

## New Variables for Contacting a Recruiter and/or Taking the ASVAB

Variable Name	Variable Description	YATS Questions
<b>DRECR2</b>	Binary: 1 = Contacted a recruiter but has not taken the ASVAB, 0 = Otherwise	Q628 is "1," has contacted a recruiter, Q645 is "1," has taken the ASVAB, and Q647 is "2," taken the ASVAB at the MEPS
<b>DASVAB</b>	Binary: 1 = Has not contacted a recruiter but has taken the ASVAB, 0 = Otherwise	Q628 is "1," has contacted a recruiter, and Q645 is not "1," has not taken the ASVAB
<b>DAPPMEP</b>	Binary: 1 = Contacted a recruiter and taken the ASVAB at the MEPS, 0 = Otherwise	Q628 is "1," has contacted a recruiter, Q645 is "1," has taken the ASVAB, and Q647 is "2," taken the ASVAB at the MEPS
<b>DAPPMET</b>	Binary: 1 = Contacted a recruiter and taken the ASVAB at a Mobile Examining Team Sight (METS)	Q628 is "1," has contacted a recruiter, Q645 is "1," has taken the ASVAB, and Q647 is "3," taken the ASVAB "somewhere else"
<b>DAPPHS</b>	Binary: 1 = Contacted a recruiter and taken the ASVAB at High School, 0 = Otherwise	Q628 is "1," has contacted a recruiter, Q645 is "1," has taken the ASVAB, and Q647 is "1," taken the ASVAB at High School

## New Variables for Schooling

Variable Name	Description	YATS Questions
<b>D16I</b>	Binary: 1 = Have not completed high school and age 16, 0 = Otherwise	Q403 is "16" and Q404 is less than "12" and Q406 equals "." (no degree)
<b>D17I</b>	Binary: 1 = Have not completed high school and age 17, 0 = Otherwise	Q403 is "17" and Q404 is less than "12" and Q406 equals "." (no degree)
<b>D18I</b>	Binary: 1 = Have not completed high school and age 18, 0 = Otherwise	Q403 is "18" and Q404 is less than "12" and Q406 equals "." (no degree)
<b>D19O</b>	Binary: 1 = Have completed high school and age 19, 0 = Otherwise	Q403 is "19" and Q404 is greater than "11" and Q406 does not equal "." (has a degree)
<b>D20O</b>	Binary: 1 = Have completed high school and age 20, 0 = Otherwise	Q403 is "20" and Q404 is greater than "11" and Q406 does not equal "." (has a degree)
<b>D1618G</b>	Binary: 1 = Have completed high school and age 16 to 18, 0 = Otherwise	Q403 is "16," "17," or "18" and Q404 is greater than "11" and Q406 does not equal "." (has some type of degree)

<b>D1618NG</b>	Binary: 1 = Have completed high school but received no diploma of any type, and age 16 to 18, 0 = Otherwise	Q403 is "16," "17," or "18" and Q404 is greater than "11" and Q406 equals "." (no degree)
<b>D1921I</b>	Binary: 1 = Have not completed high school and have received no diploma of any type and age 19 to 21, 0 = Otherwise	Q403 is "19," "20," or "21" and Q404 is less than "12" and Q406 is equal "." (has no degree)

### Other Explanatory Variable Definitions

Variable Name	Description	YATS Questions
<b>DMILCOL</b>	Binary: 1 = Definite intention for college and responded military, 0 = Otherwise	Q514 is "1," definitely going to college, and Q517 is "5," planning to service in the military
<b>Q700</b>	Grades: Values 1 to 7	Q700 is "1" for mostly A's, "2" for mostly A's & B's, "3" for mostly B's, "4" for mostly B's and C's, "5" for mostly C's, "6" for mostly C's and D's, and "7" for mostly D's and F's
<b>Q700x2</b>	Value of Q700 squared	Value of Q700 * Value of Q700
<b>Q713F</b>	Father's education: Values 7 to 20	Q713F is "7" for less than 8th grade, "8" for 8th grade, "9" for 9th grade, "10" for 10th grade, "11" for 11th grade, "12" for 12th grade, "13" for 1st year college, "14" for 2nd year college, "15" for 3rd year college, "16" for 4th year college, "17" for 5th year college/1st year graduate school, "18" for 2nd year graduate school, "19" for 3rd year graduate school, "20" for more than 3 years graduate school
<b>Q713Fx2</b>	Value of Q713F squared	Value of Q713F * Value of Q713F
<b>DQ709P</b>	Binary: 1 = Taken or plan to take physics, 0 = Otherwise	Q709 is equal to "1," have taken, or "2," plan to take physics

### Additional Binary Variables for Multinomial Logit

Variable Name	Description	YATS Questions
<b>DQ708P</b>	Binary: 1 = Taken or plan to take calculus, 0 = Otherwise	Q708 is equal to "1," have taken, or "2," plan to take calculus

<b>DA</b>	Binary: 1 = Received mostly A's in high school, 0 = Otherwise	Q700 is equal to "1"
<b>DAB</b>	Binary: 1 = Received mostly A's and B's in high school, 0 = Otherwise	Q700 is equal to "2"
<b>DB</b>	Binary: 1 = Received mostly B's in high school, 0 = Otherwise	Q700 is equal to "3"
<b>DBC</b>	Binary: 1 = Received mostly B's and C's in high school, 0 = Otherwise	Q700 is equal to "4"
<b>DC</b>	Binary: 1 = Received mostly C's in high school, 0 = Otherwise	Q700 is equal to "5"
<b>DDF</b>	Binary: 1 = Received mostly D's or F's in high school, 0 = Otherwise	Q700 is equal to "6" or "7"
<b>DFGR12</b>	Binary: 1 = Father did not complete high school, 0 = Otherwise	Q713F is less than "12"
<b>DFGR3</b>	Binary: 1 = Father has a high school diploma, 0 = Otherwise	Q713F is equal to "12"
<b>DFGR4</b>	Binary: 1 = Father has some college, 0 = Otherwise	Q713F is greater than "12," but less than "16"
<b>DFGR5</b>	Binary: 1 = Father has a college degree or greater, 0 = Otherwise	Q713F is equal to or greater than "16"



## **APPENDIX D. YATS QUESTIONS INCLUDED IN ANALYSIS**

## YATS QUESTIONS INCLUDED IN ANALYSIS

	<u>Questionnaire Section A -- Education and Employment Items</u>	
<b>Q402</b>	<p>What is the gender of the person on the line?</p> <p>1 = MALE 2 = FEMALE</p>	
<b>Q403</b>	<p>Just to be sure that the information we got earlier is correct, what was your AGE on your last birthday?</p>	
<b>Q404</b>	<p>Now I have a few questions about your educational experiences and plans. What is the highest grade or year of school or college that you have <u>completed and gotten credit for</u>?</p> <p>07 = LESS THAN 8th GRADE 08 = 8th GRADE 09 = 9th GRADE 10 = 10th GRADE 11 = 11th GRADE 12 = 12th GRADE 13 = 1st YEAR COLLEGE/JR. OR COMM. COL./VOC., BUS., OR TRADE SCHOOL (FR) 14 = 2nd YEAR COLLEGE/FR. OR COMM. COL./VOC., BUS., OR TRADE SCHOOL (SO)</p>	<p>1984 1985</p>
<b>Q404</b>	<p>07 - 12 SAME AS 84 &amp; 85 <u>4-YEAR COLLEGE OR UNIVERSITY</u> 13 = 1st (FRESHMAN) YEAR 14 = 2nd (SOPHOMORE) YEAR <u>JUNIOR OR COMMUNITY COLLEGE</u> 21 = 1st YEAR 22 = 2nd YEAR <u>VOCATIONAL, BUSINESS, OR TRADE SCHOOL</u> 23 = 1st YEAR 24 = 2nd YEAR 25 = MORE THAN 2 YEARS <u>SPECIAL CODES</u> 26 = Initial code 14 response resolved and confirmed (i.e., respondent had completed 2nd or sophomore year at a 4-year college or university but had not taken any course work beyond the sophomore level in a college degree program.)</p>	<p>1986 1987 1988</p>

Q406	<p>(Is that/Do you have) a regular high school diploma, a GED, an ABE or some other kind of certificate (of high school completion)?</p> <p>1 = REGULAR HIGH SCHOOL DIPLOMA  2 = ABE (ADULT BASIC EDUCATION) CERTIFICATE (e.g., CORRESPONDENCE, NIGHT SCHOOL)  3 = GED (GENERAL EDUCATIONAL DEVELOPMENT) EQUIVALENCY CERTIFICATE  4 = SOME OTHER KIND OF CERTIFICATE OF HIGH SCHOOL EQUIVALENCY  5 = NONE OF THE ABOVE</p>	
Q407	<p>(In October, will you be/Are you) enrolled in any school, college, vocational or technical program, apprenticeship, or job training course?</p> <p>01 = YES  02 = NO  08 = DK  09 = RE  98 = DK  99 = RE</p>	
Q416	<p>Are you currently employed, either full-time or part-time?</p> <p>01 = YES  02 = NO  09 = RE  99 = RE</p>	
Q417	<p>Are you looking for work now?</p> <p>01 = YES  02 = NO  09 = RE  99 = RE</p>	
Q436	<p>How easy or difficult is it for someone your age to get a full-time job in your community? Is it...</p> <p>01 = almost impossible,  02 = very difficult,  03 = somewhat difficult, or  04 = not difficult at all?  08 = DK  98 = DK  99 = RE</p>	
	<p>Now let's talk about your plans for the next few years. What do you think you might be doing? [PROBE: Anything else?] [ENTER CODE FOR ALL MENTIONS.]</p>	

<b>Q438A</b>	<u>First Mentioned Response</u> 01 = GOING TO SCHOOL 02 = WORKING 03 = DOING NOTHING 04 = OTHER 05 = JOINING THE (MILITARY/SERVICE) 08 = DK 98 = DK 99 = RE	
<b>Q438B</b>	<u>Second Mentioned Response</u> 00 = NO (2nd/3rd/4th) OR (subsequent/5th) mention. 01 = GOING TO SCHOOL 02 = WORKING 03 = DOING NOTHING 04 = OTHER 05 = JOINING THE (MILITARY/SERVICE) 08 = DK 98 = DK 99 = RE	
<b>Q438C</b>	<u>Third Mentioned Response</u> Codes and response alternatives are identical to Q438B.	
<b>Q438D</b>	<u>Fourth Mentioned Response</u> Codes and response alternatives are identical to Q438B.	
<b>Q438E</b>	<u>Fifth Mentioned Response</u> Codes and response alternatives are identical to Q438B.	
	<u>Questionnaire Sections B and C – Active Duty and Reserve Component Items</u>	
	Now, I'm going to read you a list of several things which young (men/women) your age might do in the <u>next few years</u> . For each one I read, please tell me how likely it is that you will be doing that.	
<b>Q503</b>	How likely is it that you will be serving in the military? Would you say...  01 = definitely, 02 = probably, 03 = probably not, or 04 = definitely not? 08 = DK 98 = DK	

<b>Q505</b>	<p>How likely is it that you will be serving in the <u>National Guard</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q507</b>	<p>How likely is it that you will be serving in the <u>Reserves</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q509</b>	<p>How likely is it that you will be serving on active duty in the <u>Coast Guard</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q510</b>	<p>How likely is it that you will be serving on active duty in the <u>Army</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q511</b>	<p>How likely is it that you will be serving on active duty in the <u>Air Force</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	

<b>Q512</b>	<p>How likely is it that you will be serving on active duty in the <u>Marine Corps</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q513</b>	<p>How likely is it that you will be serving on active duty in the <u>Navy</u>? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q514</b>	<p>Now, how likely is it that you will be going to college? (Would you say...</p> <p>01 = definitely,  02 = probably,  03 = probably not, or  04 = definitely not?)  08 = DK  98 = DK</p>	
<b>Q517</b>	<p>We've talked about several things you might be doing in the next few years. Taking everything into consideration, what are you <u>most likely</u> to be doing in (October 1985/6/7/8/9) -- that is, a year from this fall/after you finish high school)?</p> <p>01 = GOING TO SCHOOL FULL-TIME  02 = GOING TO SCHOOL PART-TIME  03 = WORKING FULL-TIME  04 = WORKING PART-TIME  05 = SERVING IN THE MILITARY  06 = BEING A FULL-TIME HOMEMAKER  07 = OTHER  08 = DK  09 = RE  98 = DK  99 = RE</p>	

<b>Q522</b>	<p>Now, I'd like to ask you in another way about the likelihood of your serving in the military. Think of a scale from zero to ten, with ten standing for the very highest likelihood of serving and zero standing for the very lowest likelihood of serving. How likely is it that you will be serving in the military in the next few years?</p> <p>RANGE: 00 (Lowest likelihood) -- 10 (Highest likelihood) 99 = RE</p>	
<b>Q628</b>	<p>Have you ever talked with any military recruiter to get information about the military?</p> <p>01 = YES 02 = NO 08 = DK 09 = RE 98 = DK 99 = RE</p>	
<b>Q645</b>	<p>Have you ever taken the three-hour written test called the ASVAB that is required to enter the military?</p> <p>01 = YES 02 = NO 08 = DK 09 = RE 98 = DK 99 = RE</p>	
<b>Q647</b>	<p>Where did you take this written test? Did you take the ASVAB...</p> <p>01 = at your high school, 02 = at a Military Entrance Processing Station (MEPS), 03 = somewhere else? 08 = DK 98 = DK</p>	
<b>Q700</b>	<p>What grades (do/did) you usually get in high school?</p> <p>01 = Mostly A's (A numerical average of 90-100) 02 = Mostly A's and B's (85-89) 03 = Mostly B's (80-84) 04 = Mostly B's and C's (75-79) 05 = Mostly C's (70-74) 06 = Mostly C's and D's (65-69) 07 = Mostly D's and F's (64 and below) 08 = DK 09 = RE 98 = DK 99 = RE</p>	

	Now I have a list of high school mathematics and technical courses. As I read each one, please tell me whether you have taken or plan to take that course in regular high school.	
<b>Q702</b>	<p>Elementary algebra (ALGEBRA I)</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q703</b>	<p>Plane geometry</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q705</b>	<p>Computer Science</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q706</b>	<p>Intermediate algebra (ALGEBRA II)</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q707</b>	<p>Trigonometry</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	

<b>Q708</b>	<p>Calculus</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q709</b>	<p>Physics</p> <p>01 = TAKEN  02 = PLAN TO TAKE  03 = NOT TAKEN  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>Q713F</b>	QUESTIONS <i>Q713F</i> named <i>Q713A</i> in 1986 through 1989.	
<b>Q713F</b>	<p>What is the highest grade or year of school or college that your father completed?</p> <p>07 = LESS THAN 8th GRADE  08 = 8th GRADE  09 = 9th GRADE  10 = 10th GRADE  11 = 11th GRADE  12 = 12th GRADE  13 = 1st YEAR COLLEGE/JR. OR COMM.  COL./VOC., BUS., OR TRADE SCHOOL (FR)  14 = 2nd YEAR COLLEGE/JR. OR COMM.  COL./VOC., BUS., OR TRADE SCHOOL (SO)  15 = 3rd YEAR OF 4-YEAR COLLEGE (JR)  16 = 4th YEAR OF 4-YEAR COLLEGE (SR)  17 = 5th YEAR COLLEGE/1st YEAR GRAD. OR  PROF. SCHOOL  18 = 2nd YEAR GRADUATE OR PROFESSIONAL  SCHOOL  19 = 3rd YEAR GRADUATE OR PROFESSIONAL  SCHOOL  20 = MORE THAN 3 YEARS GRADUATE OR  PROFESSIONAL SCHOOL  95 = MR  98 = DK  99 = RE</p>	

<b>Q714</b>	<p>Just to be sure we are representing all groups in our survey, please tell me whether you consider yourself... [IF "HISPANIC" PROBE: Do you consider your race to be white, black, Asian, or American Indian?]</p> <p>01 = white?  02 = black?  03 = Asian or Pacific Islander? (INCLUDES JAPANESE, FILIPINO, KOREAN, VIETNAMESE, PACIFIC ISLANDER, ASIAN INDIAN, OR OTHER ASIAN)  04 = American Indian or Alaskan Native?  98 = DK  99 = RE</p>	
<b>Q715</b>	<p>Are you of Hispanic background? [INCLUDES SPANISH-AMERICAN, MEXICAN-AMERICAN, PUERTO RICAN, CHICANO, CUBAN-AMERICAN, ETC.]</p> <p>01 = YES, HISPANIC BACKGROUND  02 = NO, NOT HISPANIC BACKGROUND  98 = DK  99 = RE</p>	
	<b><u>Constructed Variables</u></b>	
<b>CPYATS82</b>	<p>COMPOSITE ACTIVE PROPENSITY [Most positive response to the four Service-specific propensity (for Active Duty) questions (i.e., MINIMUM VALUE OF Q510-13)]</p> <p>01 = Definitely  02 = Probably  03 = Probably not  04 = Definitely not  08 = DK  09 = RE  98 = DK  99 = RE</p>	
<b>RSVNG84</b>	<p>COMPOSITE RESERVE/GUARD PROPENSITY [Most positive response to the four Service-specific propensity (for National Guard/Reserve duty) questions (i.e., MINIMUM VALUE OF Q505, Q057)]</p> <p>01 = Definitely  02 = Probably  03 = Probably not  04 = Definitely not  08 = DK  09 = RE  98 = DK  99 = RE</p>	

**APPENDIX E. HIGH QUALITY ENLISTMENT CONTRACTS  
MODEL**

## HIGH QUALITY ENLISTMENT CONTRACTS MODEL

An aggregate model of enlistments contracts was also developed under this research effort. This model predicts the number of high quality enlistment contracts by month for each service. High quality enlistment contracts may be defined as either AFQT category I through IIIa males with a high school diploma or AFQT category I through IIIb males with a high school diploma. This model begins by predicting the application rates for each AFQT category I through IIIb for each service. The estimation period for the model is FY84 through FY 90. These application rates are then used to determine the number of applicants for each AFQT category. Finally, historical data are used to determine the number of enlistment contracts resulting from the applicant by service.

### Data Requirements

Variables included in the estimation of the high quality application rates for each service are provided in Table E1. All applicant data used in the estimation were obtained from DMDC.

Table E1. Variable Definitions

Variable Name	Description
RECR	Number of Service specific production recruiters
UNEMP	Monthly unemployment rate for 16+ year-old population
WAGES	Relative Military to Civilian Pay Ratio
ENDSTR	Service specific enlisted FY endstrength
QTR1	Binary variable for months Oct., Nov., & Dec.
QTR2	Binary variable for months Jan., Feb., & Mar.
QTR3	Binary variable for months Apr., May., & Jun.
QTR4	Binary variable for months Jul., Aug., & Sept
DDEP	Binary variable for change in pay longevity for DEP
REZERO	Binary variable for period of no Air Force contracts

Numbers of applicants for each service by month for time period FY84 through FY90 are included in the estimations. The number of applicants for each quality group (I, II, IIIa, and IIIb) was determined from Military Entrance Processing Stations (MEPS) applicant records for each service. Records for MEPS applicants often occur more than once in the historical MEPS files as applicants reapply for entry into the Military, retake the ASVAB test or retake the physical examination. In this analysis, each applicant was made unique, ignoring duplicates if the individual applied more than once. The only exception was that an individual who made a subsequent application 24 months or more after the previous application was considered a new applicant.

To determine application rates for each quality group, the number of applicants for each group was divided by the civilian non-institutionalized population of 17 to 21 year-old males with a high school diploma and not in college. This population series was obtained from DMDC. The population series was divided by 1000 before determining the application rate for each estimation group.

The variable **RECR** is the number of service specific production recruiters by quarter. The number of production recruiters by service was obtained from DMDC. Monthly unemployment rates (**UNEMP**) for the population at least 16 years of age were obtained from the Bureau of Labor Statistics. The relative military to civilian wage (**WAGES**) was calculated as the ratio of military to civilian pay over the first four years of the recruit's military service. Civilian wages were calculated for monthly private non-agricultural wage and hours worked per week data obtained from the Bureau of Labor Statistics. Military pay included basic pay, basic allowance for quarters (**BAQ**), basic allowance for subsistence (**BAS**), tax allowances, and promotion opportunities over four years of active duty service. Service specific enlisted force endstrength (**ENDSTR**) numbers were also obtained from DMDC and included in the estimations.

Several binary variables are also included in the estimations. The time period for application is represented as categorical variables (**QTR1**, **QTR3**, and **QTR4**) with **QTR2** being a component of the constant term. **QTR1** represents the first quarter of the fiscal year, **QTR2** the second quarter, etc. Another binary variable, **DDEP**, was also included in the estimation.

This variable accounts for the change in policy which no longer allowed time in the Delayed Enlistment Program (DEP) to count towards longevity pay. This change occurred in June 1985.

In the estimations of Air Force application rates, an additional binary variable was added to the estimation equations. **REZERO** is a binary variable included in the Air Force estimations to account for the three month time period from November 1989 to January 1990. During this time period Air Force recruiters were not permitted to sign contracts with recruits due to the large force drawdowns required to meet end-of-fiscal-year force level requirements.

### **Army Estimation Results**

The results of the estimations for applications for AFQT quality groups I, II, IIIa, and IIIb for the Army are presented in Table E2. R-square values for the equations are acceptable, ranging from 0.5758 for category IIIb's to 0.6441 for category IIIa's. Root mean square errors were calculated for each equation as a measure of in-sample predictive credibility.

Differences can be seen in the determinants of application rates between AFQT groups. For example, production recruiters were statistically significant in the equations for categories I, II, and IIIa, but not for IIIb. The unemployment rate was statistically significant in all equations except for the category I equation. Wages were significant at the 90% level of confidence or greater in all of the equations. Endstrength was statistically significant in all equations except for the category IIIb equation.

### **Navy Estimation Results**

The results of the estimations for applications for AFQT quality groups I, II, IIIa, and IIIb for the Navy are presented in Table E3. R-square values for the equations are acceptable, ranging from 0.4675 for category II's to 0.6170 for category IIIb's. Root mean square errors were again calculated for each equation as a measure of in-sample predictive credibility.

Differences can be seen in the determinants of application rates between AFQT groups. For example, production recruiters were statistically significant in the equations for categories I, II, and IIIa, but not for IIIb. The unemployment rate was statistically significant in all

Table E2. Army Estimation Results

Variable	AFQT I		AFQT II		AFQT IIIa		AFQT IIIb	
	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.
RECR	0.000060	3.708	0.000273	2.660	0.000148	1.831	0.000147	1.105
UNEMP	-0.004143	-0.448	0.220173	3.772	0.244585	5.315	0.248143	3.277
WAGES	0.519514	1.653	4.215182	2.128	4.758520	3.047	9.160458	3.565
ENDSTR	0.000001	1.895	0.000005	2.496	0.000003	1.533	0.000011	4.108
DDEP	0.014249	0.783	0.061955	0.540	0.038740	0.428	-0.068764	-0.462
QTR1	-0.000076	-0.004	-0.083205	-0.737	-0.081909	-0.920	0.246128	1.680
QTR3	-0.000858	-0.070	-0.086644	-1.114	-0.071893	-1.172	-0.001259	-0.012
QTR4	-0.023834	-1.437	-0.353663	-3.383	-0.257553	-3.125	-0.179746	-1.325
Constant	-1.145555	-2.270	-9.733168	-3.060	-8.713125	-3.474	-19.501050	-4.725
R-square	0.5976		0.5699		0.6441		0.5758	
RMSE	0.0323		0.2034		0.1603		0.2638	

Table E3. Navy Estimation Results

Variable	AFQT I		AFQT II		AFQT IIIa		AFQT IIIb	
	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.
RECR	0.000039	3.352	0.000175	2.814	0.000076	2.196	-0.000006	-0.123
UNEMP	0.017471	2.319	0.178835	4.385	0.126052	5.586	0.129458	3.879
WAGES	0.133258	0.480	1.912240	1.273	2.840484	3.418	4.648066	3.780
ENDSTR	0.000003	3.507	0.000001	2.792	0.000005	2.295	0.000006	1.856
DDEP	-0.011264	-0.718	-0.138710	-1.633	-0.079350	-1.689	0.011095	0.160
QTR1	-0.030021	-1.892	-0.117785	-1.371	0.014185	0.298	0.008984	0.128
QTR3	-0.014371	-1.429	-0.086395	-1.587	-0.012522	-0.416	-0.045445	-1.020
QTR4	-0.029441	-2.182	-0.214089	-2.931	-0.071597	-1.772	-0.114925	-1.922
Constant	-1.548329	-3.434	-8.534720	-3.496	-6.564754	-4.860	-8.861632	-4.435
R-square	0.4938		0.4675		0.5646		0.6170	
RMSE	0.0269		0.1455		0.0805		0.1190	

equations. Wages were statistically significant in only the IIIa and IIIb equations. Endstrength was also statistically significant in all equations.

### **Air Force Estimation Results**

The results of the estimations for applications for AFQT quality groups I, II, IIIa, and IIIb for the Air Force are presented in Table E4. R-square values for the equations are acceptable, ranging from 0.3414 for category I's to 0.7138 for category IIIb's. Root mean square errors were again calculated for each equation as a measure of in-sample predictive credibility.

Differences can be seen in the determinants of application rates between AFQT groups. For example, production recruiters were not statistically significant in any of the equations. The unemployment rate was statistically significant in the all equations, with the exception of category I's. Wages were statistically significant in only the IIIb equation. Endstrength was also statistically significant in all equations. The binary variable **REZERO** was statistically significant in both the II and IIIa equations.

### **Marine Estimation Results**

The results of the estimations for applications for AFQT quality groups I, II, IIIa, and IIIb for the Marine Corps are presented in Table E5. R-square values for the equations are acceptable, ranging from 0.4101 for category IIIb's to 0.5849 for category I's. Root mean square errors were again calculated for each equation as a measure of in-sample predictive credibility.

Once again, some differences can be seen in the determinants of application rates between AFQT groups. However, in the Marine Corps equations, most of the variation in the dependent variable appears to be captured by the binary variables for the FY quarters. With the exception of **DDEP**, the other independent variables specified in the equation are statistically insignificant for all quality groups. The change in pay longevity calculation (**DDEP**), was statistically significant only in the category IIIb equation.

Table E4. Air Force Estimation Results

Variable	AFQT I		AFQT II		AFQT IIIa		AFQT IIIb	
	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.
RECR	-0.000035	-0.594	-0.000378	-0.813	-0.000081	-0.272	-0.000519	-1.250
UNEMP	0.003817	0.659	0.098596	2.149	0.060491	2.065	0.173058	4.221
WAGES	0.145030	0.784	1.385454	0.945	1.157119	1.236	3.083919	2.355
ENDSTR	0.000001	2.959	0.000005	3.414	0.000001	3.152	0.000006	4.178
DDEP	-0.007324	-0.699	-0.128138	-1.543	-0.069507	-1.311	-0.044278	-0.597
REZERO	0.019107	1.447	0.212126	2.047	0.143711	2.152	0.128137	1.371
QTR1	0.003058	0.288	0.005644	0.067	-0.023863	-0.445	0.131275	1.748
QTR3	0.004229	0.606	-0.006004	-0.109	-0.041068	-1.164	-0.017293	-0.350
QTR4	-0.008564	-0.907	-0.133520	-1.786	-0.098790	-2.069	-0.038987	-0.584
Constant	-0.328747	-1.382	-3.352285	-1.779	-2.574000	-2.140	-6.476245	-3.846
R-square	0.3414		0.6172		0.6116		0.7138	
RMSE	0.0184		0.1459		0.0931		0.1304	

Table E5. Marine Corps Estimation Results

Variable	AFQT I		AFQT II		AFQT IIIa		AFQT IIIb	
	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.	Coefficient	T-stat.
RECR	-0.000013	-1.960	-0.000073	-1.035	-0.000011	-0.190	-0.000006	-0.076
UNEMP	-0.002945	-1.410	0.013031	0.579	-0.002729	-0.148	0.003250	0.133
WAGES	0.125017	1.404	0.956432	0.997	0.291802	0.372	1.033457	0.995
ENDSTR	0.000005	0.400	-0.000008	-0.632	0.000002	0.161	0.000011	0.837
DDEP	0.006598	1.380	0.025706	0.499	0.032955	0.782	-0.124363	-2.229
QTR1	-0.005235	-1.104	-0.081376	-1.591	-0.085315	-2.040	-0.067391	-1.218
QTR3	-0.003461	-1.146	-0.107398	-3.299	-0.107459	-4.035	-0.112838	-3.202
QTR4	-0.011950	-2.902	-0.171782	-3.869	-0.146454	-4.033	-0.138074	-2.873
Constant	-0.137884	-0.794	0.925351	0.494	-0.137787	-0.090	-2.573552	-1.270
R-square	0.5849		0.4319		0.4295		0.4101	
RMSE	0.0078		0.0842		0.0689		0.0912	

## Predicting High Quality Enlistment Contracts

The number of high quality enlistment contracts may be predicted for each service using the estimated application rate equations presented above. Using these equations, the number of high quality enlistment contracts signed in a particular month may be predicted for each service. The number of enlistment contracts by service was obtained from DMDC.

To predict the number of high quality contracts for each month, the number of applicants by quality group must be calculated from the application rates. By specifying the values for each of the variables in the application rate equations, the application rates for each service for each quality group may be obtained. Next the population of 17 to 21 year-olds (divided by 1000) is multiplied times the resulting application rates. This will then provide the number of applicants for each quality group (I, II, IIIa, and IIIb) for each service.

Once the number of applicants by quality group has been determined, the numbers by quality group should be summed together. To determine the number of enlistment contracts for *AFQT categories I - IIIa*, the number of applicants in groups I, II, and IIIa should be summed together by service. The number of enlistment contracts may also be determined for *AFQT categories I - IIIb* by summing together by service the number of applicants in groups I, II, IIIa and IIIb.

The number of enlistment contracts may now be computed from the number of applicants using one of two equations. The equations were developed assuming a simple relationship between applicants and the number of contracts signed in a given month for each of the Services. The following equations would be used to determine the number of enlistment contracts for *AFQT categories I - IIIa*:

$$\text{Army Contracts} = 6928.128 + (0.474051 * \text{Applicants}) \quad (17)$$

$$\text{Navy Contracts} = 7809.688 + (0.017922 * \text{Applicants}) \quad (18)$$

$$\text{AirForce Contracts} = 2362.414 + (0.736558 * \text{Applicants}) \quad (19)$$

$$\text{Marine Contracts} = 3941.675 - (0.215501 * \text{Applicants}) \quad (20)$$

*Applicants* are the total number of applicants from AFQT categories I through IIIa for each service in these equations. *Contracts* are the total number of contracts for each Service predicted to be signed during the month given the number of high quality applicants.

To determine the number of enlistments contacts for *AFQT categories I - IIIb*, the following equations for each service would be used:

$$\text{Army Contracts} = 6804.242 + (0.314992 * \text{Applicants}) \quad (21)$$

$$\text{Navy Contracts} = 7543.230 + (0.054398 * \text{Applicants}) \quad (22)$$

$$\text{AirForce Contracts} = 2305.729 + (0.546245 * \text{Applicants}) \quad (23)$$

$$\text{Marine Contracts} = 3914.522 - (0.124690 * \text{Applicants}) \quad (24)$$

*Applicants* are the total number of applicants from AFQT categories I through IIIb for each Service in these equations. *Contracts* are the total number of contracts for each Service predicted to be signed during the month given the number of high quality applicants.